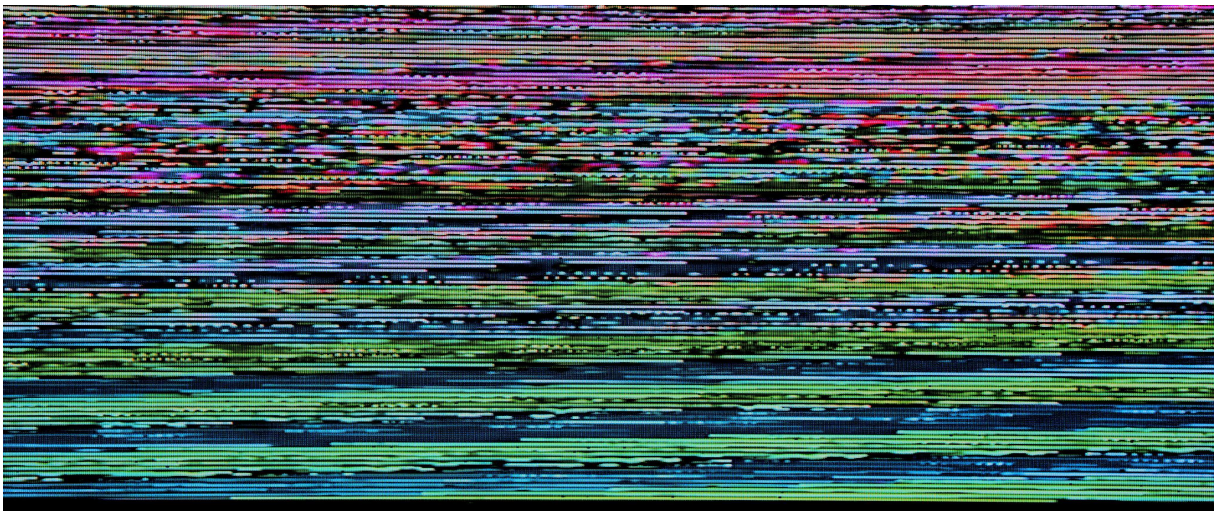
UNIVERSITY OF GOTHENBURG

DEPARTMENT OF
APPLIED IT

COMPLEXITY & RANDOMNESS

Exploring the Limits of Pattern Perception



Beppe Rådvik

Joel Pettersson

Thesis:	15 hp
Program:	Cognitive Science
Level:	First Cycle
Year:	2022
Supervisor:	William Hedley Thompson
Examiner:	Claes Strannegård
Report nr:	2022:105

Abstract

Pattern perception is a core part of human cognition, however, our capacity to process patterns is limited. If a pattern is too complex to process, we no longer perceive it as a pattern but rather as noise, thus we hypothesize that there is a limit to human pattern perception that can be measured in terms of the complexity of the pattern. To test this, we apply Aksentijevic-Gibson complexity (AG), a sophisticated measurement for the perception of complexity developed by Aksentijevic et al. [*Symmetry*, *12*(6), 948 (2020)], in a sequential comparison procedure task to test at what level of complexity participants can't differentiate between two randomly generated patterns. The result confirmed that an increase in the level of AG resulted in a decrease in performance and a significant difference between the lowest and highest level of AG. However, the hypothesized sharp decline in performance at the limits of our pattern perception was not present in the data. Despite not being able to answer the research question in detail, valuable insight into the implementation of AG in studies on visual perception was gained, as well as general insights on the implementation and application of AG and other measurements of complexity. By applying quantitative measures like AG in a wide variety of experiments, it would be possible to start outlining a general overview of the limits of human pattern perception and information processing in terms of complexity. This can provide an empirical basis for perception and information processing theories. In addition, since our perception shapes our assumptions about the world, improving our understanding of how we perceive complexity could have implications for what scientific inquiries we deem interesting and relevant and how we go about answering these questions.

Keywords

pattern perception; complexity; randomness; Aksentijevic-Gibson complexity; Visual short-term memory;

Komplexitet och slump - Kognitiva begränsningar för mönsterigenkänning

Om ett mönster är för komplext för oss att uppfatta och bearbeta så uppfattar vi det inte längre som ett mönster utan istället som brus. Vi hypotiserar att det finns en gräns för mänsklig mönsterigenkänning som kan mätas i termer av komplexiteten i mönstret. För att mäta komplexitet används måttet Aksentijevic-Gibson complexity (AG), framtaget av Aksentijevic et al. [*Symmetry*, *12*(6), 948 (2020)], som är ett mått utvecklat specifikt för att mäta mänsklig perception av komplexitet. Deltagarnas förmåga att särskilja slumpmässigt genererade mönster av olika nivåer av komplexitet testas i en sekventiell jämförelseprocedur i syfte att identifiera ett tröskelvärde där deltagarna inte längre lyckas särskilja mönstren. Analys av datan visade en signifikant skillnad mellan den lägsta och den högsta nivån av AG. Datan gav dock inte stöd för hypotesen att ett tydligt tröskelvärde skulle kunna identifieras. Även om inte forskningsfrågan kunde besvaras i detalj, så bidrar denna studie med värdefulla lärdomar om hur AG kan implementeras i studier av visuell perception samt mer generella lärdomar om implementering och tillämpning av AG och andra kvantitativa mått av komplexitet. Vi argumenterar också för att tillämpning av kvantitativa mått av komplexitet, som AG, i många olika typer av experiment gällande perception skulle bidra till utökad kunskap om begränsningar hos människans förmåga att uppfatta och bearbeta komplexa mönster, inte minst då datan skulle kunna bidra till att lägga en empirisk grund för teorier om mönsterigenkänning och informationsprocessering.

Nyckelord

mönsterigenkänning; komplexitet; slump; Aksentijevic-Gibson complexity; visuellt korttidsminne;

Foreword

The algorithm for the generation of the images used in the experiment was written by Joel Pettersson while the masking was done by Beppe Rådvik. Beppe had the main responsibility of establishing the research question and the theoretical background. Joel wrote the method and results section and the rest of the work was a collaborative effort.

Acknowledgement

We want to thank the following for their contributions:

William Hedley Thompson for being a supportive supervisor and providing valuable feedback throughout the entire process.

Gustaf Lindblad and Alexander Almér for aiding us in developing the research question.

Vicky Westermark for providing valuable feedback on the experiment design.

Aleksandar Aksentijevic & Keith Gibson for their comprehensive research on complexity.

Table of contents

1 Background **1**

1.1 Aksentijevic–Gibson Complexity in detail 2

1.2 Gestalt principles in relation to AG complexity 4

1.3 Chunking as a way to reduce complexity 5

1.4 The value of studying how we perceive complexity 6

1.5 Contribution and Hypothesis - The limits of pattern perception 7

1.6 Sequential comparisons & Visual short-term memory 8

2 Method **9**

2.1 Participants 9

2.2 Image generation 9

2.3 Design 9

2.4 Procedure 10

2.5 Tuning the difficulty 11

2.6 Ethical considerations 11

3 Results **12**

4 Discussion **14**

4.1 Interpreting the result 14

4.2 Evaluating the experiment design 14

4.2.1 Experiment design - Flaws and improvements 15

4.2.1.1 Capturing the threshold & levels of difficulty 15

4.2.1.2 Resolution & Range 15

4.2.1.3 The use of strategy 16

4.2.1.4 Degree of difference and average brightness 16

4.2.1.5 Controlled environment 17

4.3 AG applied in other studies 17

4.4 Critique of AG 19

4.5 Striving for simplicity 22

4.6 Suggestions for future research 23

5 Conclusion **25**

6 References **26**

Appendix 1 - Image generation script **28**

Appendix 2 - The study **29**

1 Background

Pattern recognition is a core part of cognition. It is central to many cognitive functions and crucial for survival, so much so that humans see patterns and regularities where there is none (Foster & Kokko, 2009). In order to effectively study pattern perception, it is essential to be able to define and differentiate between patterns and non-patterns. The absence of patterns is commonly referred to as randomness or noise. However, as argued by Aksentijevic (2015), randomness is a complicated and highly abstract concept that creates both practical and philosophical problems. Aksentijevic argues that true randomness is an elusive ideal that is difficult for humans to conceptualize and that true randomness can't be achieved. Aksentijevic also highlights the discrepancy between formal mathematical definitions of randomness and humans' subjective notion of randomness and the problems this discrepancy causes for research. One approach to studying humans' intuitive notion of randomness is by studying our ability to identify and generate random sequences. This is often done by comparing man-made sequences to sequences generated by random processes, like dice rolls or coin flips (Aksentijevic, 2015). These comparisons have highlighted that humans have a bias towards over-variation as compared to true randomness. In other words, humans perceive sequences with a lot of variation as more random than sequences with little variation, even though both sequences have been generated by the same random process (Aksentijevic, 2015). In order to bridge the gap between the abstract mathematical concept of randomness with the subjective experience of randomness, Aksentijevic (2015) suggests that we should avoid the elusive ideal of randomness and instead conceptualize randomness as an "idealized upper boundary of complexity" (Aksentijevic, 2015, p. 13).

Complexity is often defined as "the state of having many parts and being difficult to understand or find an answer to" (Cambridge Dictionary, 2022). However, Aksentijevic & Gibson (2012) proposes to define complexity as change. They argue that the more change, i.e., variation, a pattern contains, the more complex it is. If a pattern is perceived as random, it is simply too complex for the perceiver to understand and predict. Thus, whether a stimulus is perceived as random or not depends on the perceiver's cognitive capacities. When a pattern becomes too complex for the perceiver to process, it is no longer perceived as a pattern but instead perceived as randomness or noise.

Based on the notion that complexity equals change, Aksentijevic & Gibson has developed a quantitative complexity measure called Aksentijevic Gibson complexity, or AG for short. The measure was designed to quantitatively measure how costly a pattern is to process and is thus highly relevant for the study of perception (however, Aksentijevic et al. (2020) also present examples of how AG can be applied in other fields such as physics).

1.1 Aksentijevic–Gibson Complexity in detail

Aksentijevic–Gibson complexity, or AG, is a universal complexity measure based on psychological principles (Aksentijevic et al., 2020). It is designed to measure how costly a pattern is to process. Aksentijevic et al. argue that the amount of change, or variation, within a string is an important aspect of how costly it is to process, both for living and non-living agents. The higher the complexity, the higher the processing cost. For example, a string without change, like "00000", can easily be compressed by an algorithm and relatively easily memorized by a human. Thus AG measures the arrangement of changes within the structure of a string or array. This means that the measure focuses on the relationships between the elements rather than the elements themselves. For example, the string "00110" has a higher AG-value than the string "11111"; this is because the first string contains two changes (from "0" to "1" and from "1" to "0"), whereas the second string does not contain any changes.

Since AG only considers the relationships of the elements, an inversion of a string has the same level of complexity as its non-inverted counterpart. The same goes for the string's mirror image and its inversion. Thus all the following strings have the same levels of AG complexity; "00011101", "11100010" (inverted), 10111000 (mirrored), and 01000111 (mirrored and inverted), as shown in Fig 1.

		Mirrored
	00011101	10111000
Inverted	11100010	01000111

Fig 1. Table showing how all four strings are equal in terms of change. Changes are marked by curved gray arrows.

Symmetrical patterns are less complex than non-symmetrical patterns because symmetry entails repetition. Thus, the string “010” has a lower AG-value, i.e., it is less complex than the string “001”. This might seem counterintuitive as the string “010” contains two changes, whereas the string “001” only contains one. However, AG takes symmetry into account by measuring changes at higher structural levels, meaning it measures a string multiple times by analyzing parts of varying lengths in a hierarchical fashion (see Fig 2).

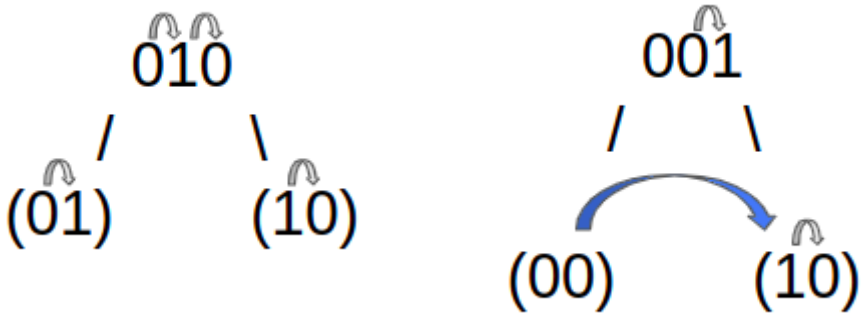


Fig 2. Diagram showing how AG accounts for change at multiple structural levels. Changes at the level of symbols are marked by curved gray arrows. Changes at the level of pairs are marked by a larger blue arrow.

The algorithm starts by identifying changes at the level of symbols. For example, in the case of the string “010”, there are two changes at this level (from “0” to “1” and “1” to “0”). The string “001”, on the other hand, only has one change at the level of symbols (from “0” to “1”); see Fig 2 and Fig 3.

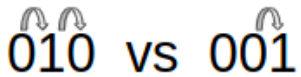


Fig 3. Comparing the number of changes at the level of symbols between the two strings.

However, when scanning the next level of structure (i.e., taking pairs of symbols into account instead of a single symbol), the string “010” does not contain any change since “01” and “10” are identical in terms of change (as is illustrated in Fig 2 and 3 by both pairs having the same number of gray arrows), and can in this case also be conceived of as mirrored repetition. The string “001”, on the other hand, does contain changes at the level of pairs since “00” and “01” differ in terms of change (as illustrated in Fig 2 and Fig 4 in terms of the difference in the number of gray arrows - “0” as opposed to “1”).

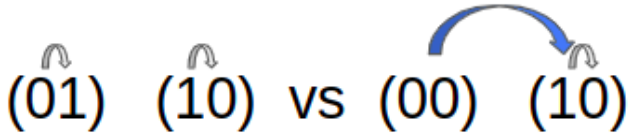


Fig 4. Comparing changes between the two strings at the level of pairs. Changes at the level of symbols are marked by curved gray arrows. Changes at the level of pairs are marked by a larger blue arrow.

Since the string “001” contains changes at two levels of structure (symbols and pairs of symbols), it measures as more complex than the string “010”, which only contains changes at the first level of structure. This highlights how “... *symmetry can be defined in terms of absence of change at higher structural levels*” (Aksentijevic et al., 2020, p. 6). If the strings are longer than three characters, the algorithm will continue analyzing change at higher structural levels. First looking at groups of three, then groups of four, etc. In summary, the more change contained within a structure, the higher the level of AG.

1.2 Gestalt principles in relation to AG complexity

To develop a sophisticated measurement of complexity that can consider all of the features human observers perceive in complexity, Aksentijevic et al. (2020) argue that we need to go beyond the information-theoretical concept of complexity that only applies quantitative measures and thus is insensitive to the structural aspect of complexity. Structure in complexity, as defined by Aksentijevic et al., is "the arrangement of a fixed number of distinct elements in a space" (Aksentijevic et al., 2020, p. 4). In other words, AG takes the order of the elements in a sequence into account (as explained earlier) rather than just the amount and type of elements contained within the sequence. To take this into account Aksentijevic and Gibson draw on the Gestalt psychologist's conceptualization of simplicity where the structure of the stimulus is of central relevance, and this major contribution to psychology therefore adds a lot to the understanding of the structural aspect of the opposite of simplicity - complexity. TheGestalt psychologists theorized that humans' perceptual framework is governed by a few simple rules, the "Minimum principle", with the purpose of striving for the conservation of energy. Perception favors simple, orderly, or in Gestalt terms, "*good*" patterns as opposed to complex, disorderly, or "*poor*" patterns. According to the Gestalt principles, good patterns are simpler because they contain fewer features and are more symmetrical, which means they can be perceived with less variety under transformations such as rotation. Aksentijevic et al. argue that a complex pattern contains more variation, or change, thus taking more effort to process. This informs

their notion that complexity equals change, highlighting the link between complexity and the cost of information processing.

These notions of good and poor patterns are not very relevant for the perception of 1D patterns such as short-strings because many of the factors, such as space and symbolic meaning, that could affect the "goodness", and thus the complexity of patterns are not present in a 1D space. It is, however, highly relevant for patterns of higher dimensions, such as the 2D images generated for this study. If AG is a good objective measure of subjective complexity, it should be able to measure the perception of these laws of simplicity. Among these laws or principles, the ones of highest relevance for perceived goodness are not clear, but generally, the following are mentioned: simplicity, stability, regularity, symmetry, continuity, and unity (Encyclopædia Britannica, n.d.). By applying AG to existing studies that concern the perception of complexity, Aksentijevic et al. (2020) provide data that supports their claim that the algorithm can detect changes in stimuli that correspond to many of these principles.

It should be noted, however, that the AG algorithm was first developed to process 1D strings and then generalized to handle 2D patterns and as Aksentijevic et al. notes, there are potentially big differences in perception of complexity between 1D patterns and 2D or 3D patterns; "goodness judgments could be difficult to capture by a generalization of the 1D approach" (2020). This problem will be discussed further on.

1.3 Chunking as a way to reduce complexity

Our memory can store a limited number of *bits* (symbols, letters, or numbers), but this limitation can be bypassed to some extent by grouping, or chunking, multiple entities together into *chunks* (Miller, 1994). By chunking patterns, the perceiver effectively reduces the perceived complexity and is thus able to process more complex patterns. In order to chunk effectively, the elements in each chunk need to be able to form a meaningful whole. Thus some sequences are easier to chunk than others. However, by applying memorization strategies, the observer can intentionally create meaningful groupings of bits even if groupings might not seem obvious at first glance. Chunking can thus be either an unconscious process or a conscious strategy to reduce complexity. For example, when memorizing long strings of numbers, it is common to use mnemonic techniques, such as associating a grouping of numbers with a word or a mental image. This has been shown to greatly improve memory performance (Anders Ericsson, 2003). Chunking is thus in line with the gestalt principles of grouping, which highlights that ordering information in groups makes it easier to process.

1.4 The value of studying how we perceive complexity

Quantitative measures of complexity, like AG, could potentially contribute to mapping out the limits of human perception and information processing and help answer the question: At what levels of complexity do we stop perceiving a stimulus as a pattern and start to perceive it as randomness or noise? The scope of this question is vast and very much intertwined with many other areas of the study of perception, such as memory, chunking, and categorization. Measures like AG complexity allow for precise control over the complexity of a stimulus and can thus be used to quantitatively measure the specific threshold value where patterns start to be perceived as noise under specific conditions. This threshold will undoubtedly vary greatly in different contexts. However, applying quantitative measures like AG complexity in a wide variety of experiments could allow us to outline a general overview of the limits of human perception and information processing in different contexts, both in terms of visual and auditory perception but also in the perception of events with random outcomes such as coinflips. This approach would require extensive research, but since the AG-value of a stimulus can be computed retrospectively, as shown by Aksentijevic et al., (2020), previous research on perception could be included in the analysis.

Gaining new insight about perception in general, and more specifically how we perceive complexity, does not only increase our knowledge of human cognition, it also has broader implications. Since our perception shapes our assumptions about the world, it also informs what scientific inquiries we deem interesting and relevant and how we go about answering these questions. Improving our understanding of how we perceive complexity and randomness could thus give us new perspectives on how our initial assumptions might misguide us when constructing scientific theories and forming scientific queries. The concept of complexity is central to many scientific domains, most notably complex systems science, which is highly interdisciplinary in nature and studies systems in a wide variety of domains such as biology, sociology, physics, and economics (Krakauer et al., 2019). We argue that improving our understanding of our perceptual mechanisms relating to complexity and randomness enables us to improve our ability to accurately conceptualize and theorize about these concepts while also improving our ability to conduct scientific research in general. Standards like AG enable the comparison of complexity in various forms, such as comparing the complexity of an image to that of a soundwave or a string of letters which deepens our understanding of the concept and gives us a language to communicate about complexity across domains.

1.5 Contribution and Hypothesis - The limits of pattern perception

Since AG is a relatively new measure, few experiments have been designed with AG in mind. The aim of this study is to contribute by conducting an experiment examining at what level of AG the participants cannot differentiate between two randomly generated patterns in a sequential comparison procedure. Thus, the primary hypothesis for the study is that we will be able to identify a threshold value where patterns start to be perceived as random noise in this particular context. By analyzing the threshold value from a range of different tasks, it would be possible to start to outline a general overview of the limits of human pattern perception in terms of AG. Such a general outline could potentially give valuable insights and provide an empirical basis for perception and information processing theories.

The secondary hypothesis of the study is that we predict that the relationship between AG and differentiability has a logarithmic character until it reaches a threshold where patterns start being perceived as noise. This hypothesis is in part based on the observation by Aksentijevic et al. (2020) that perturbation of stable structures, i.e., introducing some noise to a compact or symmetrical figure, results in a jump in complexity which follows the intuition that perception of complexity is nonlinear. It is also inspired by the Weber–Fechner law from psychophysics (Forrest W. Nutter, 2010), which can be roughly summarized as; a linear increase of the psychical stimulus, often corresponds to a logarithmic increase in the perception of the stimulus. In the case of our hypothesis, a linear increase in AG would correspond to a logarithmic increase in the perceived complexity of the stimulus.

In order to generate the stimulus for the experiment, a python script that can generate random arrays at any level of AG was written. The script is available in the appendix and can easily be used to generate stimulus for further studies. The script consists of an algorithm that generates random binary matrices and a wrapper that runs the AG algorithm (Aksentijevic et al., 2020) in R within python in order to classify the level of AG of the generated matrix. The aim is that our approach of coupling a random generator with the AG algorithm can inspire new ways of generating stimuli for psychological research on how we perceive complex stimuli.

1.6 Sequential comparisons & Visual short-term memory

To test the hypothesis, we conducted an experiment that tested the participant's ability to differentiate between two randomly generated patterns in a sequential comparison procedure. Our ability to differentiate and compare complex patterns sequentially is dependent on the capacity of our *visual short-term memory* (VSTM).

The limits of VSTM in terms of the number of objects it can hold have long been established to be around 3 to 4 items (Wheeler & Treisman, 2002). In this study, however, we are interested in the limits of our memory given a single object with high complexity, which is not as well understood. Earlier research indicates that the complexity of the stimulus indeed does matter for the capacity of VSTM. As suggested by Luria et al. (2010), this is likely since the perception of complex objects "...involves the encoding and the retention of several within-dimension feature conjunctions" (Luria et al., 2010, p. 510). Stimulus with high complexity should therefore push VSTM to its limits regardless of the number of objects perceived, and a high enough complexity should push beyond this limit resulting in an inability to discriminate differences in the stimulus. Thus we can test our hypothesis using visual perception and the theory of VSTM as the basic theoretical framework.

When conducting an experiment involving a sequential comparison procedure task, the relationship between stimulus exposure time and the ability to discriminate also has to be considered. As pointed out by von Hippel & Hawkins (1994), as the exposure time increases, the participants should be able to extract more information from the stimulus, making discriminating easier. Indeed they confirmed the hypothesis that perceptual memory increased with exposure time but that the increased time led to decreasing marginal returns before or at 1 sec. They do, however, suggest that depending on the nature of the task, exposure times beyond 1 sec may increase performance. When designing the experiment, we initially set the exposure time to 1 sec based on this research, but after receiving feedback on the high difficulty of the task, we increased this to 1.5 sec with the intent of making the task feel less stressful and potentially slightly easier. The interval time between stimuli was also set to 1.5 sec to ensure the stimulus would be kept in short-term memory. Short-term visual memory shows no loss of efficiency over the first 600 mSec and then a slow loss over the first 9 sec, after which a significant loss follows (Phillips, 1974).

2 Method

2.1 Participants

The participants were 79 persons between the ages of 18 to 77, with a gender distribution of 43 males and 36 females. A link to the study was distributed on social media. All participants were self-selected, and no incentive to participate was given.

2.2 Image generation

To test how AG affects differentiability, we designed a test that had participants view a pair of generated 2D images at various levels of AG and then had them indicate whether the two images presented were identical or not. The images were generated by visualization of two-dimensional arrays of binary strings randomly generated at the desired level of AG. This was done by generating binary strings of various lengths and then running these strings through the source code of the AG algorithm (Aksentijevic et al., 2020). The length of the strings was adjusted until the desired level of AG was reached with a margin of 0.3AG. The final images were generated at the following multi-dimensional array sizes: AG2-6x6, AG4-10x10, AG6-13x13, AG8-16x16, AG10-19x19, AG12-22x22, AG14-26x26, AG16-29x29, AG18-32x32. In the generated images, 0 corresponds to a white pixel being drawn and 1 to a black. Three images were generated at every second level of AG in the range 2-18, totaling 27 images (see Fig 6). The code for this was written in Python and can be freely accessed and used (see Appendix 1). This can be used to generate 2D matrices of any level of AG by tiny adjustments in the source code and thus be used in further research.

2.3 Design

The experiment used a 9x within-groups design. The independent variable of the study was the nine levels of AG (2, 4, 6… 18) represented by the generated images. The dependent variable was a percentage score on the number of correct answers at each level, i.e., the participants' ability to differentiate between the two patterns. The exposure time of the stimulus and the time between stimulus was 1.5 sec, and the size of the images was 150x150px with a mask size of 105x150px.

2.4 Procedure

The experiment was conducted online using the software Qualtrics (Qualtrics, 2022). The participant was presented with one of the images for 1.5 sec, followed by another image at the same level of AG for 1.5 sec with a blank image presented for 1.5 sec in-between. The second image was either the same as the first or a different image but

always at the same level of AG. The task of the participant was to indicate whether the pair presented was identical or not (Fig 5). At every level one pair was identical and one was distinct and all participants were presented with all the levels of AG with one identical pair and one distinct pair at every level, totaling 16 pairs. The order of the pairs presented was randomly generated for each participant using the random balancing function in Qualtrics. To reduce the risk of participants applying a strategy of memorizing a small portion of the image and thus making the task easier, a mask was applied to the second image presented that covered half of the image at the top, right, bottom or left part of the image. The participant could not anticipate what part of the image would be covered. They were also given instructions to perceive the image as a unified whole; "*Focus your attention on perceiving the image as a whole so you can identify any part of it*".

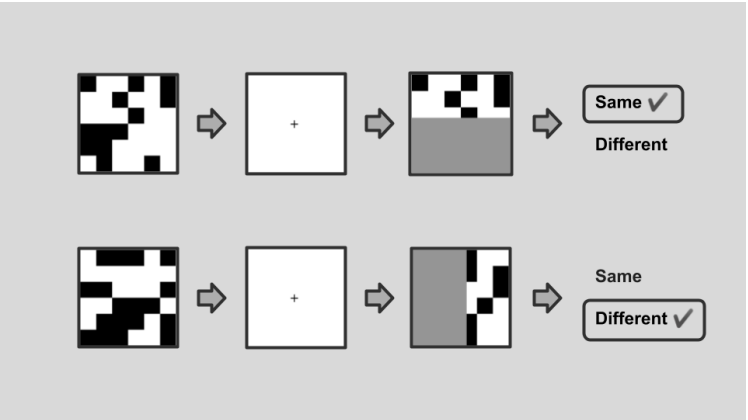


Fig 5. Illustration of the test procedure for a pair. This image was part of the instructions given to the participants.

The participants were informed that they would be presented with a series of images and that the study would test their ability to discriminate between two images (see Appendix 2). No further explanation of the purpose of the study other than this was provided. Before starting the actual test, the participants did a test run with two image pairs along with immediate feedback on their performance. The participants were advised to do the test on a computer and not a mobile device so they could properly see the image on the screen, sitting at the recommended distance from their screen and to improve concentration. The total score of the participants was gathered along with the score on each level of AG, but the participants were never informed of their results.

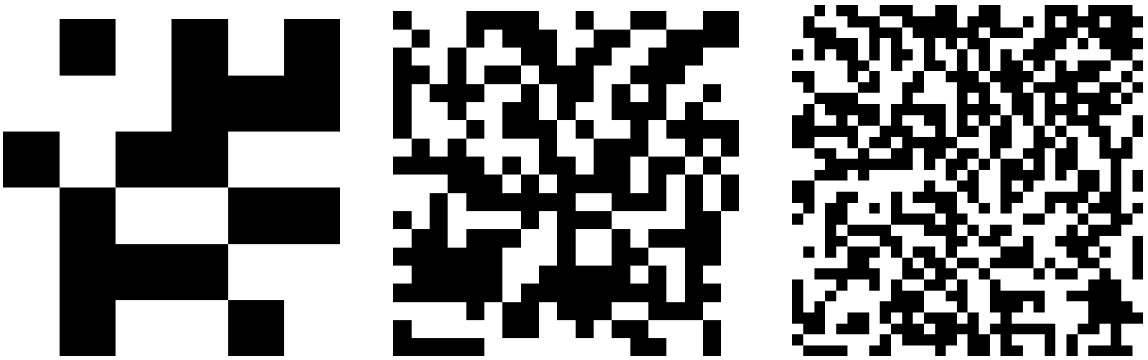


Fig 6. Image at AG2, AG10 and AG18 from left to right.

2.5 Tuning the difficulty

As no studies like this have been previously developed, we made several pilot iterations of the study to tune the difficulty of the task. Initially, the participants were presented with both images unmasked, but we observed that adopting a strategy of memorizing a small part of the image was very efficient in reducing the difficulty of the task. A mask covering 75% was implemented, but this instead made the task too difficult, so we finally settled on a mask covering 50% of the image. We also removed AG-levels 20-30 as images at these levels were too difficult to discriminate and therefore far beyond the hypothesized threshold value. Finally, exposure time, interval time, and image sizes were tuned to set the difficulty of the task at a level we deemed appropriate.

2.6 Ethical considerations

Participation in the experiment was entirely optional, and opting in was made by clicking the link distributed to the various channels. No information other than gender and age was collected from the participants to ensure complete anonymity.

3 Results

As the first step in analyzing the data, the total score for all participants was summarized to eliminate any potential outliers in participant performance. The participants scored between 8-17 in the potential range of 0-18 correct answers. The majority scored between 10-14, with a mean score of 12.15 and a standard deviation of 2.20 (Fig 7). Although there was a big span in performance, no outliers were identified in the data. A one-way within-subjects ANOVA was conducted to compare the overall effect of the level of AG on the ability to discriminate between two images. This revealed a significant effect, $F(8, 624) = 3.36$, $p = <.001$, partial $\eta 2 = .041$, meaning that more complex stimuli (higher level of AG) negatively affected ability to discriminate. Post hoc analysis with a Bonferroni adjustment revealed that ability to discriminate was statistically significantly decreased from AG2 to AG18 ($M = 0.17$, 95% CI [0.014, 0.341], $p = .021$), and from AG2 to AG14 ($M = 0.13$, 95% CI [0.019, 0.323], $p = .013$), but not between any other levels (Fig 8).

After finishing the test, the participants were asked if they adopted a strategy along with a description of the strategy, with 73% indicating that they did. The main strategy adopted was to try to memorize the pixels in the corners of the image, which means the participants actively attempted to reduce the complexity of the images by focusing on a smaller part of the image. Another common strategy was to identify large patches of black or white to get a general sense of the image's brightness. However while participants adopting a strategy tended to perform slightly better, an independent sample t-test did not reveal a significant effect in score for those adopting a strategy ($M = 12.33$, $SD = 0.28$) and those who did not ($M = 11.67$, $SD = 0.54$), $t(77) = 1.17$, $p = 0.12$. The following are some outtakes from the description of strategies by the participants:

"*1. Looked at corners (if you look at the 4 pixels in each corner, you only have 16 pixels to compare, reducing the complexity of the task, while giving you good odds) don't think i succeeded at this though.*
2. Tried to mentally "flood fill" continuous shapes."

"*Focused on the edges of the pattern to minimize the amount of information that I needed to memorize.*"

"*Don't focus on details, try to see large white areas or patterns and hope that the pattern you saw is in the part that is shown later and hope that differences are large enough to affect pattern. Also assume equal frequency for same and different.*"

"*The image was shown for too short of a time to remember the whole image. So i started focusing on the left and right edges since some part of those edges would always be visible in the second image.*"

"*I tried to note where there was a lot of white space and if I could see it again in the next image, although I feel I did not do this very successfully.*"

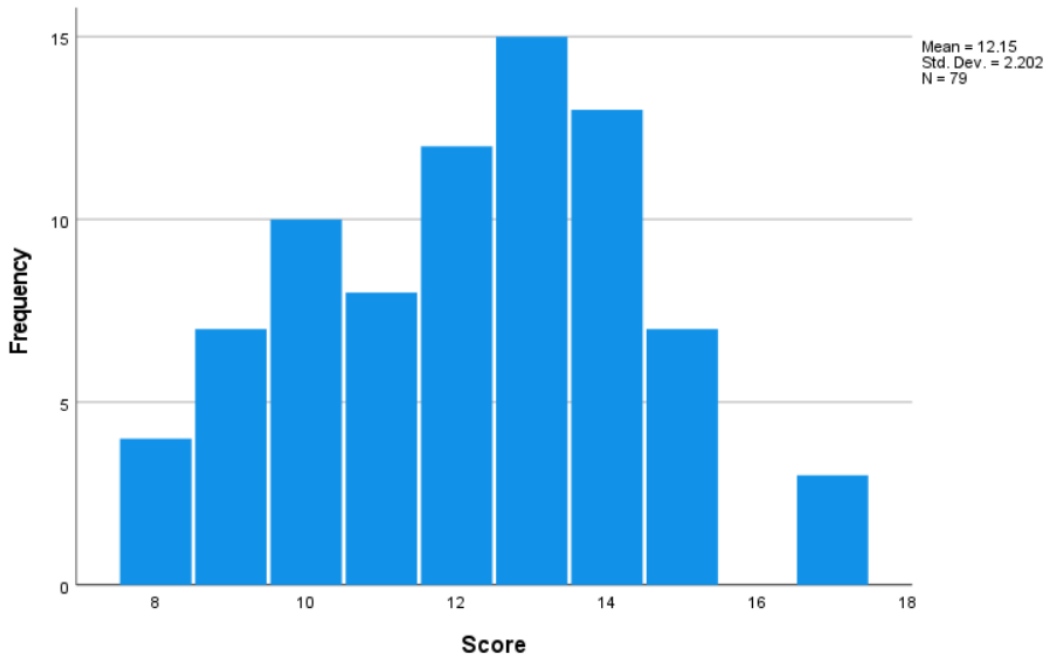


Fig 7. The participant's total score in range 0-18.

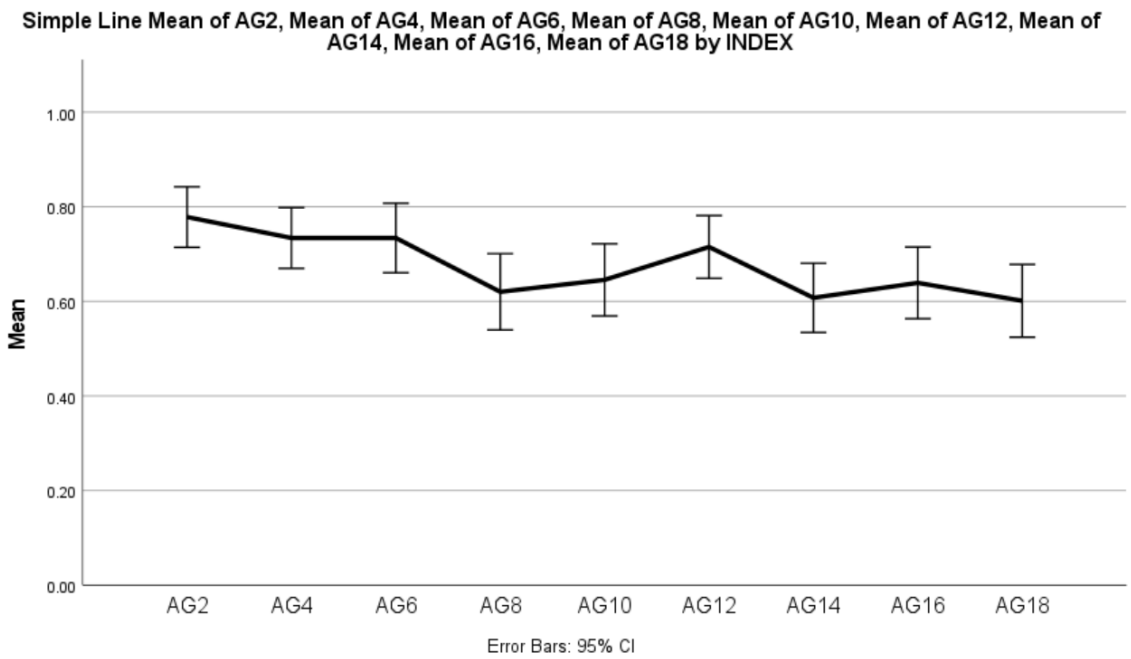


Fig 8. Simple line means of participants' ability to discriminate at AG2-AG18. The Y-axis corresponds to the participants ability to discriminate.

4 Discussion

4.1 Interpreting the result

The data shows a slow linear decrease in performance as the level of AG rises and a significant difference between the lowest and highest level of AG. However, the hypothesized sharp decline in performance was not present in the data. We anticipated that the performance across AG levels would follow a roughly logarithmic pattern and that there would be a threshold value at which point the participants would not be able to differentiate between the stimuli. Even though not statistically significant, the graph (Fig 8) suggests a potential difference between the means of AG2 through AG6 and the means of AG8 through AG18, with AG12 as an exception. The drop between AG6 and AG8 could potentially indicate the level of the hypothesized threshold, but the pairwise comparison does not yield a significant result. Thus we conclude that due to shortcomings in the experiment design, the data does not allow us to answer the research question in detail. However, we argue that improvements to the experiment design could yield data to answer the research question.

4.2 Evaluating the experiment design

As this study is the first of its kind, it is not surprising that several shortcomings in the experiment design were identified. However, this study provides a starting point for future research, and the experiment design has several advantages that could yield significant results if adequately tuned.

We have identified two main approaches to studying the perception of complexity. First, many of the studies covered in the paper by Aksentijevic et al. (2020) examine subjective reports of perceived complexity. However, subjective reports might not give us any insight into our capacities to use complex information, for example, in mental tasks such as comparisons, rotation, or alteration. Using the experiment design in this study, we contribute by testing the ability to differentiate between complex patterns, and thereby we can directly test the cost to process complexity which, as Aksentijevic et al. (2020) points out, is at the core of complexity perception and also relates to humans' cognitive capacities.

In addition, as improved measures of complexity will likely be developed in the future, using simplistic stimuli that only contain binary values increases the likelihood that these measures will be able to compute a value from our stimuli. Therefore, this study can be valuable for both existing and future studies of complexity.

4.2.1 Experiment design - Flaws and improvements

As previously mentioned, the data gathered is insufficient to answer the research question due to several shortcomings of the experiment design. In the following section, we present a number of these shortcomings, including confounding factors and suggested improvements.

4.2.1.1 Capturing the threshold & levels of difficulty

As the purpose of the study was to identify the threshold value, the task's difficulty needed to be at a suitable level to capture the threshold. If the task is too difficult or too easy, the threshold will fall outside the scope of the experiment. In order to be confident that the experiment truly captures the threshold value, the participant should be able to answer all the questions at the lower end of the range of AG correctly. The failure rate should then increase until the participant's performance is the same as chance. The result indicates that the task was too difficult as we designed the experiment assuming that most participants easily should be able to discriminate images at the lower levels of AG. However, this was not the case. Variables of the design such as exposure time, image size, and the nature of the mask all contribute to the difficulty of the task and thus need to be balanced. The exposure time, for example, should give participants enough time to perceive the patterns of the image while keeping it short enough to prevent explicit memorization strategies. This can be achieved by conducting multiple pilot studies. Once the threshold is well within the range of measured values, variables can then be altered systematically to examine their effect on the threshold value. By examining how the threshold is affected by the variables, we can start to outline the limits of perception and information processing in this particular context. As previously mentioned, applying this methodology to a range of experiments would provide an empirical foundation that can provide new insight into how humans perceive complexity.

4.2.1.2 Resolution & Range

In order to get a nuanced view of the relationship between AG complexity and differentiability in our experiment, the range of AG values and the incremental increase between each level of AG need to be set appropriately. Our experiment had a relatively low resolution as the level of AG increased by 2 for each level, and the range of values was set to a fairly wide range (2-18), especially considering many of the studies analyzed by Aksentijevic et al., (2020) fall in the range of AG1 through AG10. However, it’s worth noting that these studies examined AG in a completely different experimental context and that very little testing of AG has been conducted. In order to be able to answer the research question, the resolution and range need to be tuned to an appropriate level by conducting pilot studies. As a starting point, we suggest a range of AG1 through AG10 with an incremental increase of 1.

4.2.1.3 The use of strategy

Since the study aims to study the perception of complexity, measures were taken to reduce the use of conscious and unconscious strategies that reduce the perceived complexity of the patterns. The participants were instructed to "Focus your attention on perceiving the image as a whole so you can identify any part of it," and a mask was added to the stimuli to prevent the participants from focusing on only one part of the pattern. However, these measures were insufficient as 73% of participants reported using a strategy.

The most common strategy was to only look at parts of the picture. The use of strategies like these heavily reduces the quality of the data in at least three ways. First, as mentioned previously, the use of strategies might reduce the perceived complexity of the patterns and thus inhibits an accurate measure of perceived complexity. The use of strategies also shifts the participants' focus in a way that detracts from their initial impressions of the stimuli as a uniform object which is what we wanted them to perceive. The use of strategy could also be a distraction as the participants might focus on sticking to their strategy rather than observing the stimuli. The fact that the task is demanding and monotonous, as indicated by some participants, may also contribute to an increased adoption of strategies.

In improving this experiment it is then of high priority to mitigate the use of strategies by the participants, especially strategies involving memorizing small parts of the image which reduces the perceived complexity. A potential solution to this is to mask the image in a much less predictable way by masking the majority of the image and keeping a small viewable window of the stimulus which is moved around between each stimulus. Another way to reduce the adoption of strategies is to mix the main task of image discrimination with other tasks that breaks the training effect and confuse the participant about the nature of the task, forcing them to constantly adapt to new tasks and thus reducing the effect of any strategies adopted.

4.2.1.4 Degree of difference and average brightness

Another potential confounder that needs to be accounted for is the degree of difference between the patterns contained in the stimulus pairs. Since the patterns are randomly generated, there is a chance that the two stimuli in a pair contain a similar set of pixels resulting in a similar pattern, or that all of the pixels differ between them, resulting in a big contrast. In other words, the degree of difference between the pairs at each level of AG varies. The same goes for the average brightness of the stimuli. Some of the patterns contain more black pixels and thus have a lower average brightness, whereas some contain more white pixels and thus have a higher average brightness which could be another confounding factor.

The effect of these confounding factors could be reduced by using multiple pairs of stimuli at each level of AG. Ideally, the difference in pattern and average brightness between the two stimuli in a pair should be calculated quantitatively, which would allow for an examination of these factors as independent variables. We predict that the differences are especially relevant for lower levels of AG. This could potentially be a suitable focus for further research. Including multiple different patterns at each level of AG would also balance out eventual bias caused by the positioning of the mask, thus resulting in an overall improvement of the experiment.

4.2.1.5 Controlled environment

Finally, the experiment design would benefit from being conducted in a controlled, physical environment instead of conducting it online via a survey tool. When conducting it online, the participants may be in environments affecting their concentration or using a device that impairs their vision of the stimulus. The participants in this study were advised to use a computer for the task, but around 37% still used a mobile device. An independent sample t-test indicated that participants on mobile devices performed slightly worse ($M = 11.84$, $SD = 2.02$) than those on normal computer devices ($M = 12.34$, $SD = 2.30$) but no significant effect was revealed, $t(77) = 0.99$, $p = 0.16$. With a larger sample size, this effect might have been significant. Thus, it would be preferable to conduct the experiment in a controlled environment to reduce potential confounders like these.

4.3 AG applied in other studies

Not many studies have been designed specifically to utilize AG as a measure, and the few that have are not very relevant for this study, so they will not be discussed here. Aksentijevic et al. (2020) do however apply it retrospectively in a meta-study to test its correlation with subjective responses from several studies, which gives valuable insight into how AG performs in various contexts. The studies were picked because they are well-known and cover many different aspects of psychological complexity in tasks testing performance in auditory, visual, or speech. All of the studies have developed their own measurements for subjective perception of complexity. Without a standardized measurement like AG, they cannot be compared to each other. Aksentijevic et al. findings show that 23 out of 26 variables in the studies correlate significantly with AG and that it explains up to 60% of the behavioral variance across the studies. Thus we can see the accuracy of AG and the value of applying a standardized measure of complexity in these studies.

Concerning the research question of this particular study, little insight into the hypothesized threshold value can be gathered from analyzing the studies mentioned above or the limited amount of studies using AG (around 20 were found in total). A more

fruitful approach to provide the theoretical background to the threshold value is to gather insight from the vast amount of research on the cognitive limits of perception, such as studying theories like VSTM and chunking, which will be discussed further. However, an intriguing comparison can be made between this study and a simulation of a pattern going from order to disorder made by Aksentijevic et al. (2020) to illustrate AG's ability to capture the dynamic aspect of complexity (e.g., a biological process where something changes over time). The patterns are created in a 20x20 matrix, and for every timestep, the compact figure is being gradually eroded. From this, a complexity curve can be plotted (Fig 9).

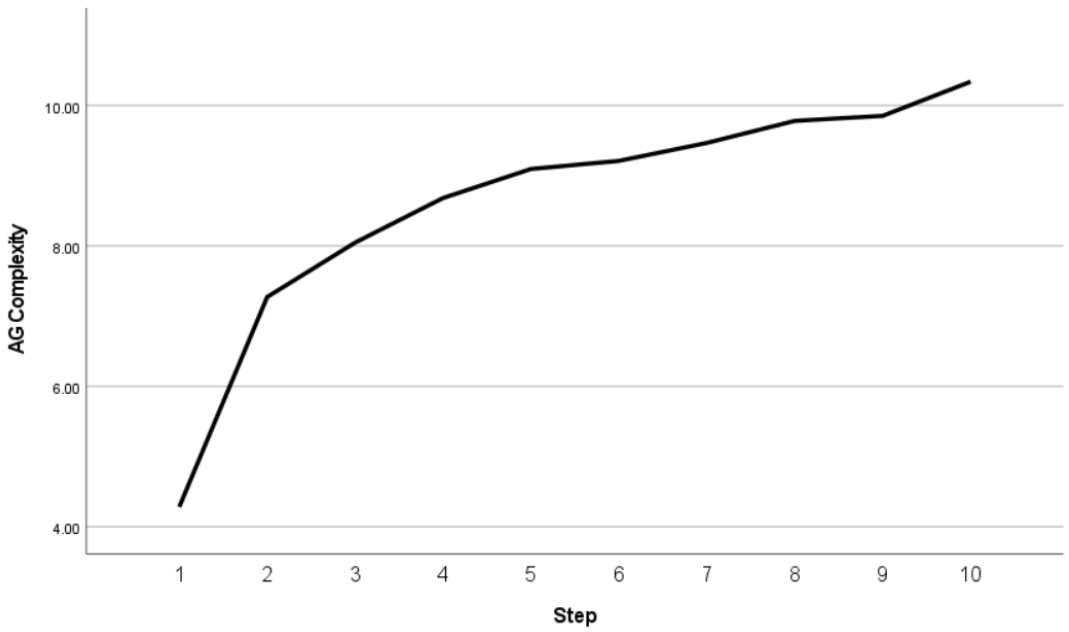


Fig 9. "Transition to disorder. 2D AG complexity over the course of the simulation. Snapshots of the dissipative process have been taken every 20 steps." (Aksentijevic et al., 2020, p. 27). *Snapshots of the figure being eroded can be seen in Fig 10.*

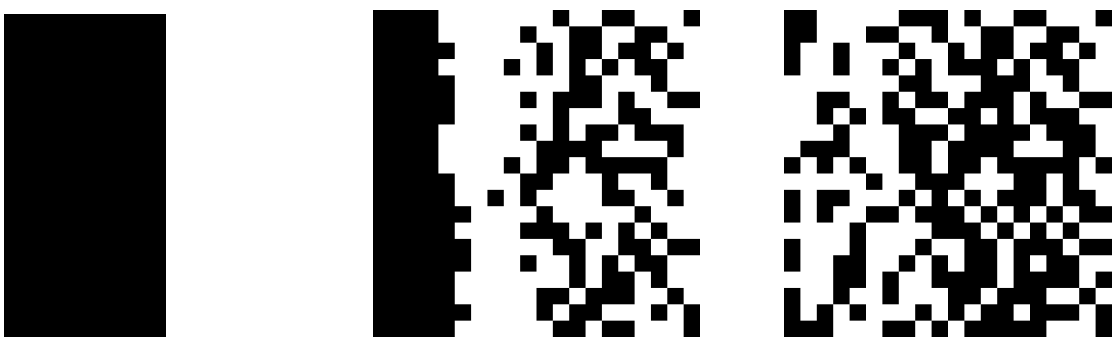


Fig 10. Transition to disorder. Snapshot from steps 20 (1), 120 (6), and 200 (10).

This graph cannot be compared to the graph plotted for the results in our study, but it can nevertheless provide valuable insight as it illustrates the perception of complexity in terms of AG. A jump in complexity at step 40 (2 in the graph) can be observed, which, as Aksentijevic et al. points out, "is related to the earlier observation that stable structures are more easily perturbed and confirms the intuition that complexity perception is nonlinear"*(Aksentijevic et al.*, 2020, p. 27). The graph is also in line with our hypothesis that complexity follows a logarithmic pattern. Note that the accuracy of this graph is dependent on the ability of AG to accurately capture the actual human perception of complexity. Ideally, it would be compared to data from subjective responses of perceived complexity, reversing the retrospective application of AG explained above.

4.4 Critique of AG

A standardized measure of complexity that can be applied across domains would be very valuable. However, that requires a measurement that is highly refined and well-tested to ensure that it is able to accurately measure and account for all relevant features of complexity regardless of domain. AG should therefore be thoroughly examined and compared with other potential measurements.

A commonly used proxy of complexity are measures of entropy, but these have been harshly criticized for their inability to capture the compression of data into patterned chunks (Gauvrit, Zenil, Delahaye, & Soler-Toscano, 2014) or as Aksentijevic et al. puts it; measures of entropy ignores the structural aspects of complexity. Another approach is algorithmic complexity which is a measurement of the length of the shortest algorithm that can output a particular string, or how long it would take the algorithm to run given a particular input. The term is primarily used in theoretical computer science to compare algorithms and improve computation time (Devopedia, 2022). This approach, like AG, can take the structure of the string into account and is thus more sophisticated than measures of entropy. However, Aksentijevic and Gibson (2012) highlight that encoding in algorithmic complexity can result in simple patterns being encoded in a complex way while, at the same time, any encodings can have many different meanings. It also does not consider the cost of processing complexity which, as they argue, is highly relevant for a measure of complexity that not only applies to computer algorithms but also to human observers. We agree with this assessment, and AG then, from our findings, is the most sophisticated measure of complexity currently conceptualized in regard to human perception.

A promising indicator of AG's generalizability is illustrated in its correlation at around 0.7 in studies using stimuli with 2D patterns (Aksentijevic et al., 2020), in particular, as no

other measures attempt to calculate the complexity of 2D patterns. However, as pointed out by Aksentijevic and Gibson themselves, the algorithm for 2D arrays is simply a generalization of the 1D algorithm. The complexity of 2D patterns is calculated by taking the average complexities of its rows, columns and diagonals which are then summarized to get a value of the complexity of the entire pattern without applying weights to any of the arrays. As indicated by Aksentijevic and Gibson themselves, there are potential issues with generalizing the 1D algorithm to a 2D scenario. As explained earlier, the concept of the Gestalt psychologist's notion of "good" patterns does not matter much for the complexity of a string, as many factors affecting the complexity of 2D and 3D patterns, such as space and symbolic meaning, do not exist in a 1D pattern. In addition, in the perception of a 1D pattern where processing is mainly sequential, periodicity plays a major role in the perception of complexity alongside symmetry, whereas with a 2D pattern parallel processing dominates, which causes symmetry to be the biggest contributor to complexity (Baylis & Driver, 1994). This claim is also supported by the Gestalt theories that symmetry is critical for a pattern to be perceived as good. Despite this potential weakness in the algorithm, however, Aksentijevic et al. findings indicate that the algorithm is able to detect changes in stimuli that correspond to many of the Gestalt principles. While AG does not explicitly quantify these principles, the most important among them being symmetry, it seems to be able to at least partially capture them by the unique approach of scanning and measuring the changes at various levels and the interaction between these levels, as explained earlier.

The potential limits of AG can, however, start to be identified in more subtle cases where our intuition of perceived complexity clashes with the calculation given by the algorithm. One such problem is illustrated in Figure 11. The image on the left is according to the measurement less complex (by 2 AG levels) than the image on the right; however, our intuition is that due to the black column in the right picture it would be easy to identify and discriminate in our task. Given an information-theoretical conceptualization of complexity, it would make sense that the image on the right is more complex as it contains more information but as Aksentijevic and Gibson claim that AG also takes structure into account, one would assume that a calculation of these patterns would result in the image on the right being less complex, or at least that they are on a similar level of complexity. AG thus seems to be overly tuned to the amount of information or entropy in the pattern in this case. The image generation developed for the task in this study could potentially generate images with higher structure, as the image on the right in Figure 11, but as mentioned previously, it could be controlled for by using multiple pairs of stimuli at each level of AG.

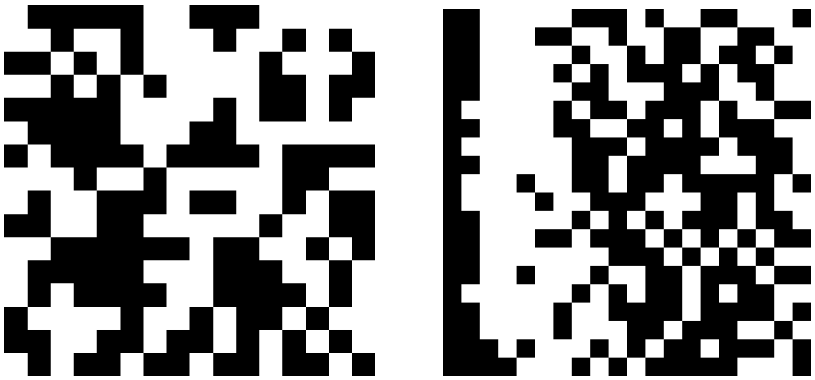


Fig 11. On the left is the image generated by our script at 16x16 (AG-value: 7.83) and on the right is an image from the Transition to disorder simulation at 20x20 (AG-value: 9.85).

Another illustration of this issue in a more extreme case can be seen in Figure 12.

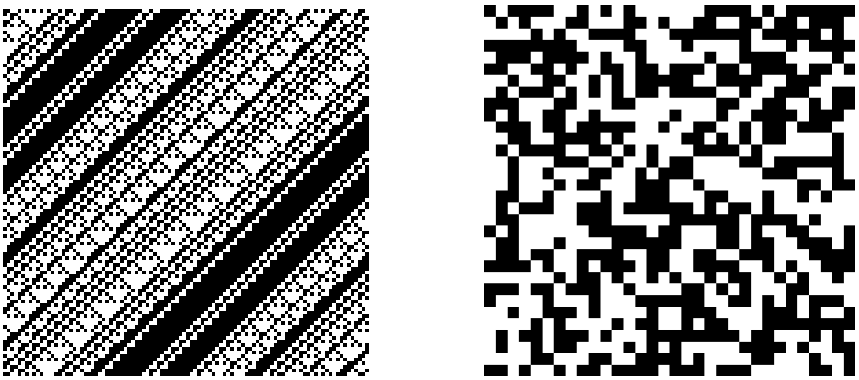


Fig 12. Image generated by elementary cellular automata (Aksentijevic et al., 2020) (AG-level: 54) on the left and image generated by our script to the right (AG-level: 18).

Admittedly the pattern on the left is of high complexity and difficult to comprehend, but it does not appear to be as random as the pattern on the right and certainly not at three times as high level of complexity. In fact, a distinct structure and repeating patterns can be identified in the image, which cannot be identified in the other image, which appears more random. Thus for the task in our study, it is probable that the participants would be able to discriminate the image generated by elementary cellular automata but not the image generated by our script, even though it is of a much higher level of AG. In comparison with other patterns generated by means of elementary cellular automata, the image in the figure is of the lowest complexity, and Aksentijevic et al. (2020) argue that the most random-looking pattern of them has the highest complexity which is in agreement with Kolmogorov's claim that randomness equals high complexity (Durand & Zvonkin, 2007). AG should, in other words, be able to account for perceived randomness,

but this analysis indicates that it might not be able to do this adequately when the patterns being compared have been generated in different ways, which is problematic for our aim of identifying a threshold value and potentially also for other research questions. This then not only highlights the issue that AG may be overly sensitive to the amount of information in a pattern but also that it is poorly adapted to compare the complexity of patterns generated with different methods and thus possibly also to compare complexity across domains. As this is central to the value of AG as a standardized measure, it should be of high priority for future research.

Finally, to truly be a standardized measure of complexity, AG would have to be able to calculate not only 2D arrays but arrays with even more dimensions. In addition, the current algorithm can only take an input of two binary strings, but there are many cases where an input of symbols other than simple binary strings would be desired. This is an absolute necessity in order to analyze inputs such as sounds and colors.

4.5 Striving for simplicity

Section 4.2.1.3 on the use of strategy, gives an account of how the participants utilized strategies to reduce the perceived complexity of the stimuli in an attempt to make the task easier. We argue that the adoption of strategies inhibits an accurate measure of perceived complexity and that focusing on sticking to a strategy potentially distracts the participants from the task. As explained earlier, we tried to prevent the adoption of strategies by implementing a mask to the stimuli; however, most of the participants still reported that they adopted a strategy. As suggestions for improvements to the experiment design, we propose adding other tasks in between rounds of the sequential comparison procedure, forcing the participants to constantly adapt to new tasks and thus reducing the effect of any strategies adopted. However, it is possible that the use of strategy cannot be controlled for when using these types of stimuli.

Although speculative, it is plausible that humans inevitably will resort to some sort of strategy, conscious or unconscious, in order to process complex information. This would be in line with the reasoning of Foster & Kokko (2009), stating that humans tend to automatically try and find patterns and regularities. We speculate that resorting to strategy, conscious or unconscious, could be conceptualized as an attempt to reduce the perceived complexity to make the information easier to process, making it possible to perceive patterns at a simpler level of abstraction. In other words, if a stimulus is of too high complexity (above the threshold value), the perceiver automatically tries to reconceptualize the input to a lower resolution. This makes sense from an evolutionary perspective as reducing the perceived complexity entails conservation of energy, which is in line with the Gestalt principles of striving for simplicity. We suspect that how, and to what extent, the information is reconceptualized and compressed is highly dependent

on context and that the perceiver will strive to find an optimal level of resolution for the task at hand. The perceiver can thus be primed by the context to perceive the stimuli at a particular level of abstraction. In the case of our experiment, some participants reduced the resolution of the perceived patterns by adopting the explicit strategy of only focusing on the corners of the image, since this was assumed to be sufficient for performing the task of differentiating between the stimuli.

We argue that reconceptualization can be both a conscious and unconscious process. A good example of this is chunking. Depending on the task at hand, people will automatically group, or chunk, certain objects together. However, chunking can also be utilized as a deliberate memorization technique (Anders Ericsson, 2003). We hypothesize that reducing perceptual complexity to make information easier to process is common in a wide variety of domains and scenarios. The concept of noise and randomness itself is an example of this, highlighting how large amounts of complex information can be reduced to just one phenomenon. One could even extend this line of reasoning to encompass other cognitive processes beyond perception. For example, in his book *Thinking Fast and Slow*, Kahneman highlights how when we struggle to find an answer to a hard question, we tend to substitute it with a related, easier question (Kahneman, 2011). This could also be conceived as an unconscious attempt to reduce complexity or at least to make information easier to process. Although speculative, we deem this line of inquiry highly relevant for cognitive science as it is relevant for many areas of cognition.

4.6 Suggestions for future research

As previously stated, we argue that in order to get a general overview of the limits of pattern perception and information processing in terms of AG, the measure needs to be applied in a multitude of experimental settings across different modalities. In this section, we suggest a number of specific experiments where AG can be applied, as well as more general areas that could benefit from a standardized measure of complexity.

For further research on visual pattern perception, we suggest creating variations of classic studies, such as (Phillips, 1974), using stimuli generated at different levels of AG. The python script used to generate the stimuli for this study can serve as a starting point for designing stimuli for these experiments. Our study examined differentiability in a sequential comparison procedure, however, differentiability could also be studied in the context of side-by-side comparisons. As this study has shown, AG is relevant to the study of visual short-term memory. As sequential comparisons require the perceiver to store the information in the visual short-term memory, our experiment would have benefited from more knowledge on how the level of AG complexity of objects affects our capacity to memorize them. This could be examined further by conducting experiments

by having participants memorize sets of symbols of varying complexities and measuring how the level of AG impacts performance. As more complex patterns are more costly to process, it's possible that our memory span, i.e., the number of objects we are able to memorize is reduced if the individual objects are of high complexity as the complexity, to some extent, inhibits effective chunking. AG could also be relatively easily applied to studies on mental rotation in a similar manner.

All experiments mentioned above, examine our capacity to process complex visual patterns in terms of using them in mental tasks such as differentiation, memorization, or rotation. In order to gain a better understanding of how we perceive complexity, the results could be compared to subjective judgments of complexity. This could provide insight into if our immediate subjective judgment of the degree of complexity within a pattern matches our ability to process the pattern in a mental task, thus highlighting the difference between patterns that we report as random and patterns that are too complex for us to process.

Since AG aims to be a universal complexity measure, it enables us to compare the limits of our different perceptual systems (vision, hearing, touch). Identifying threshold values in many different experiments concerning all different perceptual systems would give valuable insights into the limits of perception. It also enables us to compare how our pattern perceptions differ when the patterns are presented sequentially through time, for example as a video, as opposed to being presented as a static image. This difference can, to some extent, be studied for all perceptual systems. However, as previously highlighted by our critique of AG, in order to successfully compare the limits of our perceptual systems, AG needs to be further developed. Thus future research should prioritize the development of universal measures of complexity.

5 Conclusion

This study investigated at what level of AG humans cannot differentiate between two randomly generated patterns in a sequential comparison procedure. Based on the results, no definitive conclusion can be made about the hypothesized threshold value. However, the data showed that an increase in the level of AG resulted in a decrease in performance and a significant difference between the lowest and highest level of AG. Although the data was inconclusive regarding the research question in detail, valuable insight into the implementation of AG in studies on visual perception was gained, as well as general insights on the implementation and application of AG and other measurements of complexity.

This study can serve as a pilot for further research as it identifies a suitable range of AG(1-10) for this specific task. We also identified the need to control for the degree of difference in pixels, and the difference in average brightness, between the two stimuli. The fact that most participants reported that they used a strategy could potentially give valuable insight into how we handle complex stimuli, namely that humans strive to find patterns and that this could make controlling for the use of strategy impossible in this experimental setup. We also speculate that humans tend to, consciously or unconsciously, try to reduce the perceived complexity to process the information more effectively and that this could be a fruitful line of inquiry for further research.

We also highlight that AG, in some cases, is poorly adapted to compare the complexity of patterns generated with different methods and that this could be a hindrance to comparing complexity across domains. As universality is central to the value of AG as a standardized measure, we argue that improvement of the 2D algorithm should be of high priority for future research.

Finally, a general theory of complexity applied to human cognition enables insight into cognitive faculties across multiple domains, contributing to the progress of cognitive science. We propose a research approach of using a standardized measure of complexity in a wide variety of experiments to gain insights that could contribute to mapping out the limits of human cognition using AG complexity. We also argue that knowledge about perception deepens our understanding of ambiguous concepts like complexity, randomness, and noise, and this enables us to improve our ability to accurately conceptualize and theorize about these concepts across various scientific domains.

6 References

Aksentijevic, A. (2015). Randomness: off with its heads (and tails). *Mind & Society*, *16*(1–2), 1–15. https://doi.org/10.1007/s11299-015-0187-7

Aksentijevic, A., & Gibson, K. (2012). Complexity equals change. *Cognitive Systems Research*, *15–16*, 1–16. https://doi.org/10.1016/j.cogsys.2011.01.002

Aksentijevic, A., Mihailovic, A., & T. Mihailovic, D. (2020). Time for Change: Implementation of Aksentijevic-Gibson Complexity in Psychology. *Symmetry*, *12*(6), 948. https://doi.org/10.3390/sym12060948

Anders Ericsson, K. (2003). Exceptional memorizers: made, not born. *Trends in Cognitive Sciences*, *7*(6), 233–235. https://doi.org/10.1016/s1364-6613(03)00103-7

Baylis, G. C., & Driver, J. (1994). Parallel computation of symmetry but not repetition within single visual shapes. *Visual Cognition*, *1*(4), 377–400. https://doi.org/10.1080/13506289408401715

Cambridge Dictionary. *Complexity*. (2022, May 11). Retrieved May 13, 2022, from https://dictionary.cambridge.org/dictionary/english/complexity

Devopedia. *Algorithmic Complexity*. (2022, February 19). Retrieved May 19, 2022, from https://devopedia.org/algorithmic-complexity

Durand, B. & Zvonkin, A. (2007). Kolmogorov Complexity. In: Charpentier, É., Lesne, A., Nikolski, N.K. (eds) *Kolmogorov's Heritage in Mathematics*, 281-299. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-36351-4_14

Encyclopædia Britannica. *Perception*. Britannica Academic. (n.d.). Retrieved May 19, 2022, from https://academic-eb-com.ezproxy.ub.gu.se/levels/collegiate/article/perception/110465

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, *104*(2), 301–318. https://doi.org/10.1037/0033-295x.104.2.301

Forrest W. Nutter, F. W. N. (2010). *The SAGE Encyclopedia of Research Design*. In Weber-Fechner Law (Second Edition (Revised Edition) ed., pp. 1613–1615). SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n494

Foster, K. R., & Kokko, H. (2008). The evolution of superstitious and superstition-like behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1654), 31–37. https://doi.org/10.1098/rspb.2008.0981

Gauvrit, N., Zenil, H., Delahaye, J. P., & Soler-Toscano, F. (2014). Algorithmic complexity for short binary strings applied to psychology: a primer. *Behavior Research Methods*, *46*(3), 732–744. https://doi.org/10.3758/s13428-013-0416-0

Kahneman, D. (2011). *Thinking, Fast and Slow* (1st ed.). Farrar, Straus and Giroux. 97

Krakauer, D. C., Gell-Mann, M., Arrow, K., Arthur, B. W., Holland, J. H., Lewontin, R., Morrowitz, H., Flack, J. C., Dunne, J., & West, G. (2019). *Worlds Hidden in Plain Sight: The Evolving Idea of Complexity at the Santa Fe Institute, 1984–2019 (Compass)*. SFI Press.

Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2010). Visual Short-term Memory Capacity for Simple and Complex Objects. *Journal of Cognitive Neuroscience*, *22*(3), 496–512. https://doi.org/10.1162/jocn.2009.21214

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352. https://doi.org/10.1037/0033-295x.101.2.343

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290. https://doi.org/10.3758/bf03203943

Qualtrics XM. (2022, May 9). *Qualtrics*. Retrieved May 13, 2022, from https://www.qualtrics.com/

von Hippel, W., & Hawkins, C. (1994). Stimulus exposure time and perceptual memory. *Perception & Psychophysics*, *56*(5), 525–535. https://doi.org/10.3758/bf03206949

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. https://doi.org/10.1037/0096-3445.131.1.48

Appendix 1 - Image generation script

The code in “main.py” was written by us. All other files were downloaded from the package provided by the authors of *Time for Change: Implementation of Aksentijevic-Gibson Complexity in Psychology.*
https://github.com/joel-p/kandidatarbete/tree/master

Appendix 2 - The study

Intro

Hi, thank you for choosing to participate!

The purpose of this study is to examine people's ability to discriminate between two images. The study takes approximately 4-5 minutes to complete.

If you agree to participate you will be presented with a series of image-pairs and your task will be to indicate whether the images presented are the same or not. After you have completed the study you will be presented with a score based on how many correct answers you managed.

Your participation in this study is entirely anonymous, your name will never be connected to your results. The only data gathered in this study is your input at the start and end of the study and your score (% correct answers). Information that would make it possible to identify you or any other participant will never be included in any sort of report.

The study is conducted by Joel Pettersson and Beppe Rådvik as part of their bachelor thesis at the University of Gothenburg. If you have any questions about the experiment you may contact us at: guspetjocx@student.gu.se.

The experiment works best in Google Chrome (or other chromium-based browsers)! Please avoid using Firefox if possible.

Please confirm that you have read the above information and want to participate in this study.

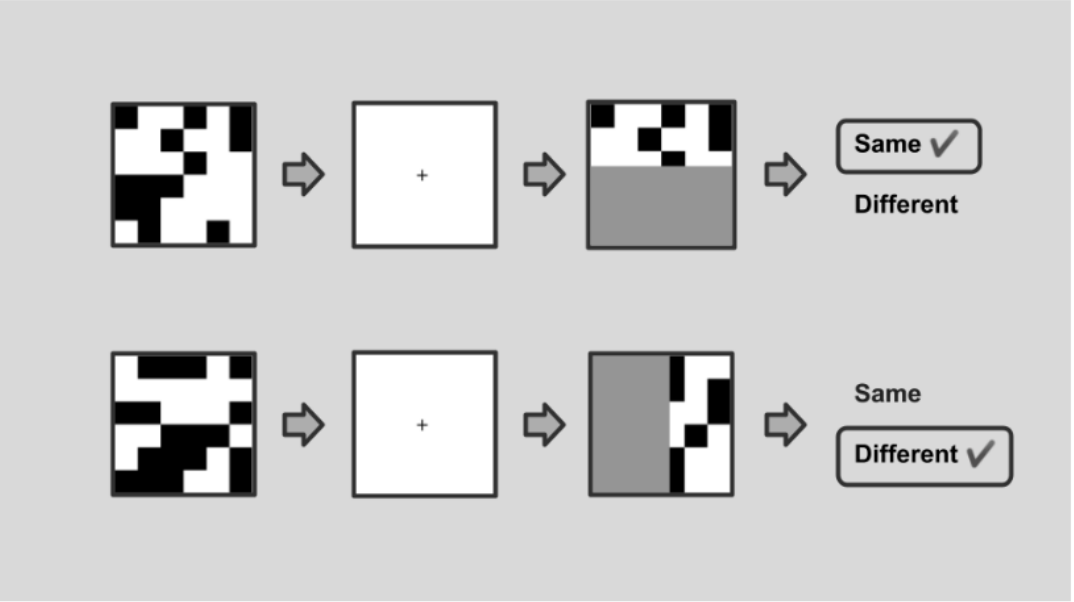
I understand the conditions and want to participate

I do not want to participate

Instructions

You will now be presented with a series of image-pairs with each sequence containing the following elements:

1. An image of a pattern
2. A blank image
3. The same pattern again but partially masked, or a new pattern that is partially masked.
4. A question where you answer whether you think the patterns were the same or different.

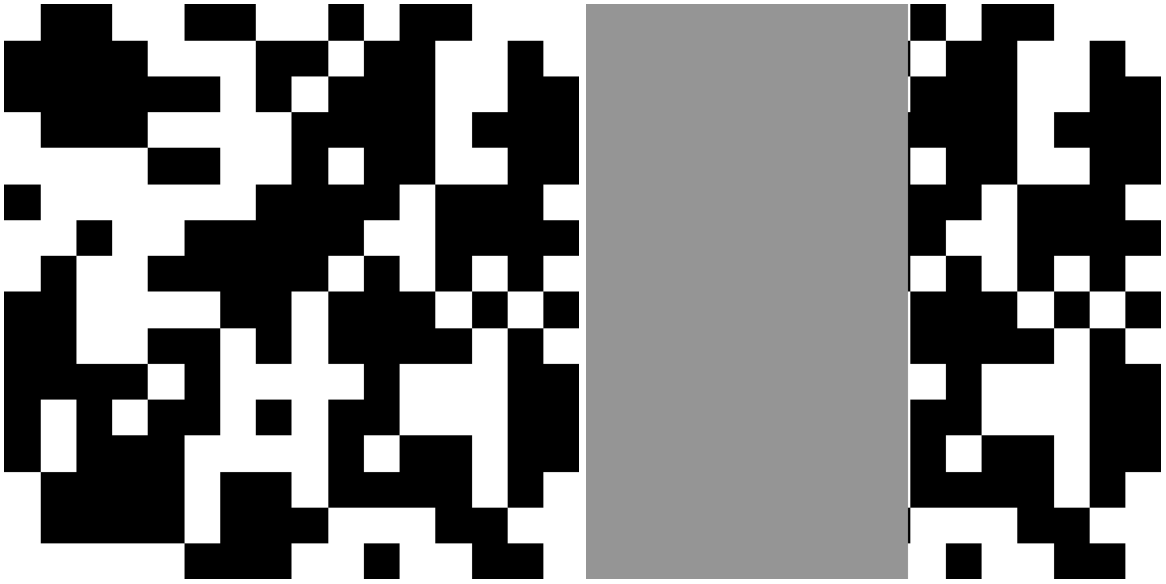


- Each image will be presented very briefly, so keep your attention on the image area.
- Answer as quickly as possible without error.
- Focus your attention on perceiving the image as a whole so you can identify any part of it.

We will start with a series of 4 pairs to get you warmed up. These will not contribute to your result. Make sure you sit at a comfortable distance to your screen and can focus on it for the next few minutes without distractions. We do NOT recommend you do this test on a phone.

AG8 - *Different*

AG8 - *Same*



Same or different?

Same

Different