02464 Artificial Intelligence and Human Cognition

Miniproject 1

Authors

Francesco Balducci - s250200
Nicola Lunardi - s251989

October 12, 2025

Contents

1 Introduction

The aim of this course is to explore the connection between artificial intelligence and human cognition. Neural networks have recently been designed to mimic brain structure, with nodes (neurons) connected by links (synapses) enabling processing.

Comparisons between humans and AI must be fair, considering their nature is fundamentally different. There doesn't exist only one definition for intelligence: for example, an LLM completing a poem is based on statistical training of word likelihood, while human intelligence can involve sensitivity and personal experience.

Human intelligence is categorized into domains, reflecting related abilities. The top-level domain, Spearman's G, explains much (but not all) intelligence across domains, followed by numerical, spatial, linguistic abilities, and elusive areas like social, physical, and emotional intelligence.

Intelligence must be understood with respect to the computing architecture. David Marr's framework divides computation into levels: computational (problem encoding), algorithmic (steps used to solve the problem), and implementation (physical form).

1.1 Memory as benchmark

Learning is core to system evolution and performance, whether artificial or natural. It involves memorization, encoding general rules or patterns rather than specific details, such as recognizing numbers by simpler shapes. Though computers and humans can be compared computationally, implementation differs: humans store "concepts" in brain regions like the hippocampus, unlike computers storing isolated information that scales independently of the nature of the memorized items.

2 Experiment setup

Having to memorize and then recall characters and subsequently analyze the results allows us to shed light on the phenomena involved in human memorization. To this end, we have created two experiments consisting of *free recall* and *serial recall* exercises. In the first case, a series of characters are shown and then the participant is asked to repeat them in random order, while in the second case, the participant is asked to re-present them in the same order in which they were shown.

Both experiments were conducted on a publicly visible site[1]. This is created ad-hoc using various libraries, where it is possible to specify the experiment settings and your student number. The site then displays instructions on how to correctly perform the experiment and shows a series of characters, asking the user to enter those they remember, giving them the option to stop when they feel they have remembered everything possible. It then downloads a CSV file with the list of characters shown and those reported, for the analysis that will follow.

[1] Free recall - Serial recall - all other code is also available in open format in the respective GitHub repositories.

In the two experiments, the following configurations were used:

2.1 Free Recall

- **Number of characters displayed**: 20 - a number of elements greater than what we assume short-term memory can store (7) is required and we want to be able to notice the difference between the recall rate in the first region (1st - 3rd element), middle (4th - 17th element) and last region (18th - 20th element)
- **Types of characters shown**: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789!@#$%& - it is necessary to have at least 20 characters to show
- **Display time**: 0.5 characters/sec
- **Character repetition**: no - random effects that could allow a greater number of items to be remembered without using short-term memory must be avoided
- **Variants**:
  - Rate: increases the time a number is displayed to 1 character/sec
  - Delay: introduces a 10-second delay between the memorization phase and the recall phase
  - Task: introduces a short-term memory task between the memorization phase and the recall phase (you must memorize and recall in reverse order an additional 5 numbers that are not counted for the purposes of general recall)
- **Expected effects**:
  - Primacy: the numbers shown first are stored in long-term memory and are therefore more easily remembered than those in the middle. The task (task variant), by not affecting short-term memory, does not change this effect. The longer display time (rate variant), by allowing more time for long-term memorization, increases this effect. The delay before recall (delay variant), by not introducing additional elements into long-term memory, does not change this effect.
  - Recency: the numbers shown last are stored in short-term memory and are therefore more easily remembered than the central ones. The task (variant "task"), by affecting short-term memory, reduces this effect. The longer display time (variant "rate"), leaving more time for long-term storage alone, does not change this effect. The delay before recall (the "delay" variant), by not introducing additional elements into short-term memory, does not change this effect.

2.2 Serial Recall

- **Number of characters displayed**: 10 - a number of elements greater than what we assume short-term memory can store (7) is required

- **Types of characters shown**: 123456789 - at least 10 characters must be shown
- **Display time**: 0.5 characters/sec
- **Character repetition**: yes - it is also necessary to evaluate how the regular repetition of elements allows a greater number of them to be remembered
- **Variants**:
  - Chunk: alters the random generation of characters by introducing regularities (format of the type A B B A and so on)
  - Articulatory suppression: instructs the user to continuously vocalize a meaningless term ("eerb") during memorization
  - Tapping: instructs the user to perform a task with their hands (draw a square with their fingers) during memorization
- **Expected effects**:
  - Pattern: there is no linear relationship between the number of items shown and those that can be remembered, as memory exploits patterns (the "chunk" variant) for efficiency reasons. This increases the number of items remembered correctly.
  - Articulatory rehearsal: we use subvocalization as a submodule of short-term memory to remember more items. Suppressing this loop ("articulatory suppression" variant) reduces the number of items remembered correctly.
  - Tapping: as proof of what was stated in the previous point, the introduction of a task that, although added, does not involve the same areas related to memory as a task with the hands (variant "tapping") does not produce the same effect. This does not change the number of items remembered correctly.

3 Analysis of results

The result analysis was performed using a Jupyter notebook created using the Pandas and MatPlotLib libraries to read and visualize the statistics that follow.

The data collected by the team members for each test were combined as if they came from a single subject (40 trials for each experiment and its variants, 320 trials in total), in order to obtain greater statistical robustness. Subsequently, the following were calculated:

- Proportion of correct answers
- 95% confidence interval using Wilson's method (better for smaller samples), to understand whether the differences in proportion are significant and not due to chance

A bar chart was generated to represent the number of correct answers for each position, allowing the primacy and recency effects to be visualized.

In addition, for the free recall experiment, three 95% confidence intervals were calculated for the proportion, respectively for the first three (primacy), the central (4-17), and last three positions (recency).

This made it possible to directly compare the intensity of the two effects and observe how they change as experimental conditions vary.

3.1 Free Recall

Baseline condition In the basic experiment (presentation rate of 0.5 s/item), subjects correctly remembered **29.5%** of the items (CI: 26.4–32.8%). Analyzing the three regions specifically, we observe the confidence intervals reported in Table 1. The bar chart (Figure 1) shows the expected trend: the first and last elements are the most frequently remembered.

Presentation rate Halving the presentation rate resulted in an increase in the proportion of correct responses (**32.3%**). However, because the confidence intervals overlap, this difference is not statistically significant (CI: 29,1-35,6%). The same applies to the confidence intervals reported in Figure 2 , as summarized in Table 1.

Delay before recall In the delayed recall experiment, the proportion of correct responses decreased to **24.9%**. Nevertheless, this difference is not considered statistically significant, since the calculated confidence interval overlaps with that of the baseline (CI: 22.0–28.0%). Similarly, the proportions of the three regions did not decrease significantly (Table 1).

Secondary task In the experiment with the secondary task, accuracy decreased significantly to **15.4%** (CI: 13.0–18.0%). The only range that showed a significant decline was the last one (decrease of approximately *46*%), as can be seen in Table 1.

Region	Baseline (0.5 s/item)	Rate (1.0 s/item)	Delay	Secondary task
First	32.5–49.8	30.1–47.3	18.8–34.3	17.4–32.6
Central	16.1–22.6	18.1–24.8	14.2–20.5	10.2–15.7
Last	57.8–74.5	69.2–84.0	51.1–68.3	13.1–27.1

Table 1: *95% confidence intervals for recall performance by region and condition.*

3.2 Serial Recall

Baseline condition In the baseline condition, subjects correctly recalled 51.1% of the items (CI 46.1–55.9%). Again, the graph (Figure 2) shows the expected trend: the first items are recalled with greater accuracy, while the last items are more prone to error.

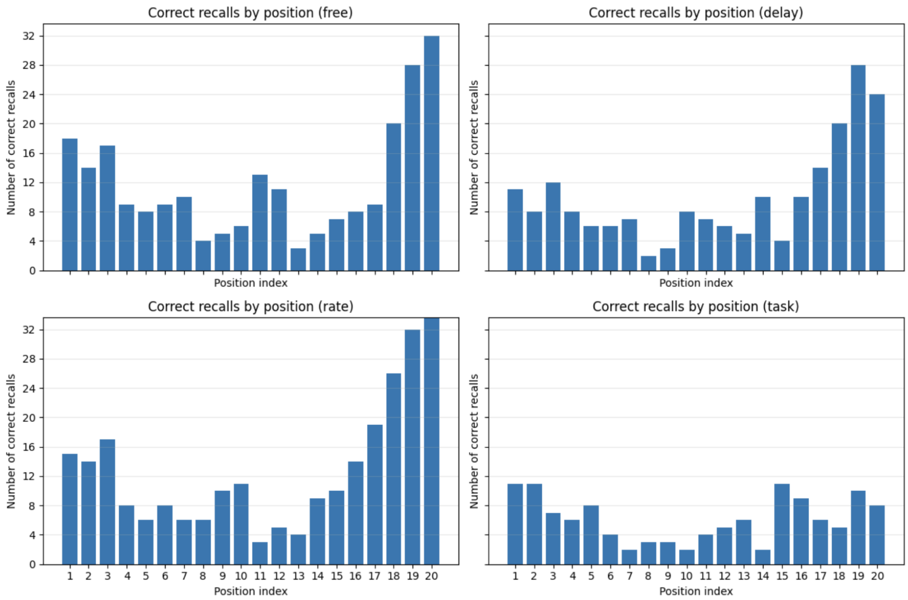

Figure 1: *Free recall performance by condition.*

Articulatory suppression With articulatory suppression, accuracy dropped to 30.8%, a significant decline as the confidence intervals do not overlap (CI: 26.4–35.4%). This shows that articulatory suppression interferes with sublocal repetition, which, in the usual condition, helps to memorize information in the short term.

Chunking condition Under the chunking condition, the proportion of correct responses increased significantly to 82.3% (CI: 78.2–85.7%). This shows that grouping the elements to be memorized reduces the capacity required to perform the task, simplifying it significantly.

Tapping condition Finally, with tapping, the proportion of correct responses was 39.8% (CI: 35.1–44.6%). Again, the decrease is considered significant, implying a negative effect on memorization (contrary to what was expected from the study)possibly because the task distracted participants and reduced their attention..

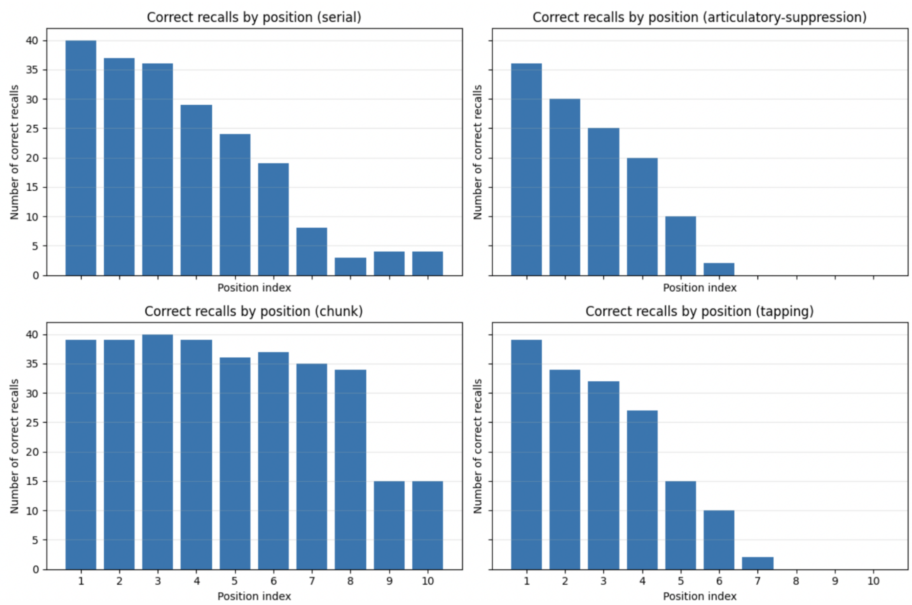

Figure 2: *Serial recall performance by condition.*

Error types The error analysis was carried out by comparing, for each test, the numbers presented with those recalled by the subjects. The most frequent errors were: 8 mistaken for 7, 3 for 2, and 2 for 3. We expected pronunciation-related errors, but participants instead tended to confuse numbers that were close to each other, indicating a difficulty in distinguishing adjacent values rather than random recall errors.

4 Conclusions

Overall, the results obtained replicate some of the effects described by the theory. In free recall, the correct-position response curve clearly showed the effects of primacy and recency. The speed of presentation and the delay at the end of the presentation were not significant. On the contrary, the secondary task led to a sharp drop in accuracy, especially in the last positions, demonstrating, as expected by the theory, the relationship between the recency effect and short-term memory. In serial recall, the conditions of articulatory suppression and chunking produced the expected results: the former reduced performance, while the latter improved it. On the other hand, finger tapping reduced performance, contrary to what we expected.

A Contributions

All participants in the assignment have compiled and/or reviewed the present document and all source codes. However, for organizational reasons, we decided to be responsible for different parts of the document, as shown in the following table for reference:

	Francesco Balducci (s250200)	**Nicola Lunardi** (s251989)
Introduction	70%	30%
Experiment setup	60%	40%
Analysis of results	30%	70%
Conclusions	40%	60%

Table 2: Table of responsibilities

B About LLM usage

Although the LLM was used to enhance students' understanding of various concepts, including website building and data analysis methodologies, it was not involved in the actual writing of this document, except for paraphrasing content originally written by hand.