The speed of information processing at memory competitions: limited by reading in short tasks and declining as a power law for longer times

Bastian Wiederhold
bastian.wiederhold@lmu.de

Faculty of Biology
LMU Munich
March 31, 2025

Abstract

We analyze information rates associated with top performances in memory competitions and reveal three phenomena. First, in tasks with short memorization time, information processing reaches up to 42 bit/s with most of the time spent on reading, suggesting that mental associations are formed even more rapidly. Second, record performances show a remarkable concordance across time scales: the processing speed depends on memorization time as a power law. Third, despite dramatic improvements in scores and mnemonic strategies over the last decades, the differences in information rates across memorization tasks remain remarkably consistent.

1 Introduction

Memory techniques have been known since antiquity [Gri22]. Starting about three decades ago, formal rules for memory competitions were established, which allowed a comparison of the competitors' control of mnemonic strategies. The most famous is the "memory palace" or "method of loci". During memorization, competitors first transform information into "mnemonic images": memorable objects, people or scenes. These images are then linked to prominent locations along a mental walk. During recall, competitors reimagine the same mental journey and decode the mnemonic images to retrieve the stored information. The databases of competitions provide us with the unique opportunity to study a controlled setting in which participants have optimized their strategies for years individually, and for decades collectively.

Recently, it has been argued that the human brain is "slow", processing information only at around 10 bits/second including examples from memory competitions [ZM24]. Our systematic analysis will reveal a more diverse picture. We analyze data from classical competitions of the International Association of Memory (IAM), the modern format Memory League (ML), and speed-memory.com. Combined, these tasks span a wide range of memorization-time periods from one second to one hour. To compare different disciplines, we calculate the minimum information processing rates in bit/s necessary to explain top performances. We find that for short tasks, information rates greatly exceed 10 bit/s and seem to be limited by reading rather than memorizing. Quite generally, the rates follow a power law as a function of the memorization time.

2 Results

2.1 Small difference between reading and memorization times in rapid tasks

Using a memory palace involves three different processes: reading, associating and navigating. For example, a competitor in the decimal digits discipline first maps the visual input, the digits,

onto specific prememorized mnemonic images. This process is comparable to standard reading as, indeed, the most popular strategy to learn the map is to assign letters to the digits $0-9$ [Artb]. The resulting images are associated to salient positions of the mental walk through short creative stories. Once a location is occupied, the competitor mentally navigates to the next location in the memory palace. Among these three processes, it seems most natural to equate "memorization" with the process of forming associations.

To understand the time spent on the three processes, we asked top competitors in ML cards, digits and words for the time they require to read the items of the tasks (Table 1). The self-reported reading times could be fairly accurate, as competitors train extensively over years to minimize total time. In ML words, reading is comparable to everyone's understanding of the term, whereas in ML cards and digits, reading refers to the aforementioned transformation into mnemonic images.

Furthermore, we calculated a realistic minimum of the reading time based on research on single-word reading, which suggests that 200 milliseconds elapse due to visual processing before the visual input becomes available to the mnemonists' minds [Hau+12]. We received values in line with the self-reported times and the estimate of 50 bit/s for general reading [PK57].

Self-reported times and the estimated minima indicate that reading requires most of the memorization time. If mental associations were formed solely during the difference Δ between the memorization time and the reading time, processing rates in memory would be significantly higher than indicated by the task itself.

	Performance (P)	P $[\frac{\text{bit}}{s}]$	Reading time (RT)	RT $[\frac{\text{bit}}{s}]$	Difference (Δ)	$\Delta\,[\frac{\text{bit}}{s}]$
Alex Mullen	ML cards 11.89 s 52/52 4	18.97	200 ms estimate $\frac{52}{2} = 26$ items $\simeq 5.2$ s	≈ 43	6.7 s	≈ 34
			Self-reported ≈ 9 s	≈ 25	2.9 s	≈ 78
Andrea Muzii	ML digits 9.75 s 80/80 [Muz]	27.25	200 ms estimate $\lceil \frac{80}{3} \rceil = 27$ items $\simeq 5.4$ s	≈ 49	4.4 s	≈ 60
			Self-reported ≈ 8.25 s	≈ 32	1.5 s	≈ 177
Don Michael Vickers	ML words 12.77 s 45/50 [Vic]	42.04	200 ms estimate 45 items $\simeq 9$ s	≈ 60	3.8 s	≈ 141
			Self-reported ≈ 9 s	≈ 60	3.8 s	≈ 141

Table 1: Differences Δ between the reading and memorization time and the associated information rates for selected memory athletes. To determine the length of the reading phase, we asked the competitors and estimated the minimum based on 200 ms per item [Hau+12]. Even these conservative estimates constitute a large fraction of the memorization time. If reading and the formation of associations were entirely sequential, Δ would be the association/memorization time corresponding to the speed in the right-most column. In ML cards the task is to memorize the order of 52 cards in a shuffled deck, in ML digits an 80-digit decimal number and in ML words a sequence of 50 words. Andrea Muzii is using a three-digit system, converting three-digit decimal numbers into one mnemonic image, so there are 27 items to be perceived [Arta]. Alex Mullen achieved the time with a system encoding two cards in one mnemonic image, so 26 items need to be perceived [Tho16].

2.2 Power law of memorization speed

Across memorization tasks, the greater the number of items to be memorized, the longer the time competitors require. The reading time per item is not expected to increase decisively, but more time is required to consolidate the information in memory until the start of the recall phase. It is, therefore, not surprising that the information rates are lower for longer tasks. Intriguingly, when we calculated the minimum information rates necessary to explain the top performances, we found a clear power law of information rate over the memorization time (Figure 1 **A**). We fitted the function aT^b for two parameters a, b to the data points provided by the official records using least squares based on relative residuals. As a function of the duration T of the

memorization phase in seconds, the resulting curve of the information rate R is

$$R(T) \approx 41.71 \cdot T^{-0.35} \ \frac{\text{bit}}{s}. \tag{2.1}$$

Note that, compared to standard forgetting power laws [WE91], Equation (2.1) does not describe the proportion of correctly recalled items over time, but rather the achieved information rates over different time spans. The dependence of the memorization time on the information of the complete task is also described by a power law rather than a linear relationship (Figure 1 **B**). This demonstrates that the power law is not a simple large-set variant of Hick's law [PS18]. As a function of the complete memorized information I in bit, the memorization time in seconds is described by

$$R'(I) \approx 19.5 \cdot 10^{-4} \cdot I^{1.57} \ \ s.$$

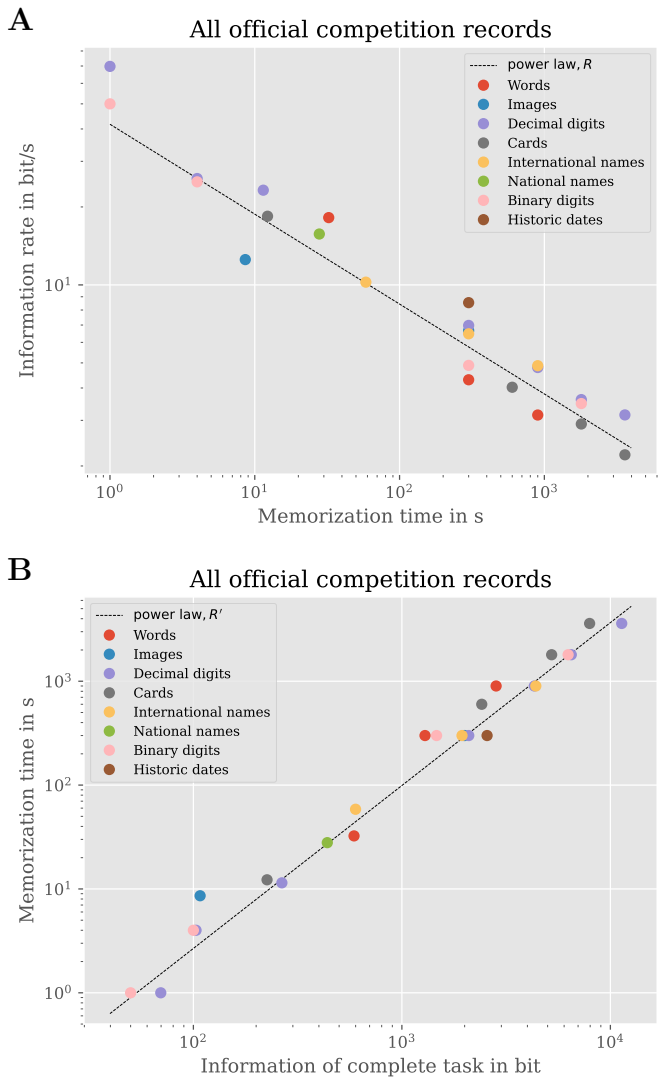

Figure 1: **A** Official world memory records: information rates in bit/s as a function of the length of the memorization time. The results of 1 s and 4 s memorization time are from speed-memory.com. The remaining results with a memorization time smaller than 60 s are ML records and the bottom right-hand cloud corresponds to the records from classical competitions. The vertical alignment of parts of the data results from the memorization at classical competitions being either $5, 15, 30$ or 60 min. An overview of the competition formats can be found in Section 4.2, information calculations in Section 4.3 and the list of records in Section 4.5. The speed-card record in classical competitions of Shijir-Erdene Bath-Enkh is 12.74 s. It is reassuring that the score is similar the ML cards record of 12.25 s, as this is the only discipline directly comparable between ML and classical competitions. **B** Official world memory records: memorization time as a function of the information of the complete tasks in bit.

Using a simple probabilistic model, we related the power law to the probability to remain error-free in tasks with a given number of items and given time (Section 4.4). The model is based on the assumption that competitors aim for error-free memorization and that long tasks can be divided into shorter tasks with independent error probabilities. The model predicts that competitors' success-rate for perfect memorization is high if enough time is given for a certain number of items, but rapidly decreases if the memorization time becomes shorter than a certain threshold. This seems to match observations on the error-rates in ML (Figure 2).

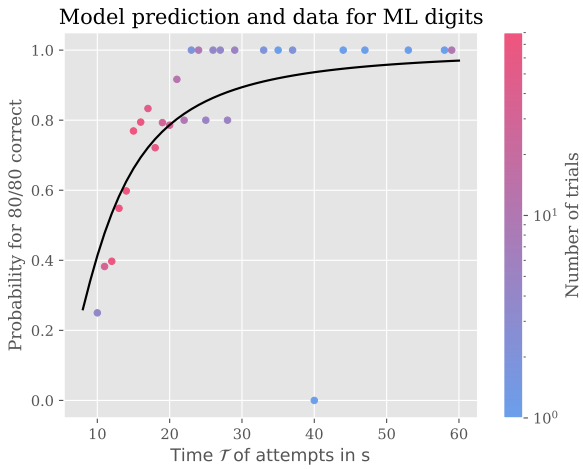

Figure 2: Empirical success probabilities for achieving 80/80 in ML digits compared to the prediction of the probabilistic model. Shown are data for the success probability of trials, binned in intervals of one second, by competitors Alex Mullen and Andrea Muzii, the official and unofficial ML record holders. Note that around 85% of trials are in the time between 10 and 20 s, and larger times are insufficiently sampled. For example, there was only a single attempt in the 40 s bin. The curve depicts the prediction $\mathbb{P}(A(80, \mathcal{T}-r))$ of the model. We used the parameters of the power law 2.1, a reading time r of 8 s (Table 1) and obtained $p = 0.72$ for the free parameter of the model using weighted least squares, with the number of attempts of a certain time as weights.

2.3 Consistent differences between encoding systems

Classical competitions comprise memorization of decimal digits, binary digits and cards (Table 2). Over recent decades, world records have improved steadily with no clear indication that the trend is weakening (Figure 3 **A**). Among the disciplines based on symbolic sets, decimal digits stand out with strikingly linear record progression (Figure 3 **B, C, D**). The comparison between the achieved information rates across symbolic sets shows consistent relations. Decimal digits are processed faster than cards: this trend is shown convincingly in Figure 3 **E** for disciplines from 10 s to 60 min memorization time over the past 30 years. The comparison between decimal and binary digits is more complicated: Figure 3 **F** shows that there is a clear and consistent gap of around 40% for the 1 s and 5 min disciplines, but essentially the same memorization speed in the 30 min discipline.

3 Discussion

A recent study revealed the conundrum of human cognition: whereas our sensory systems are able to process enormous quantities of information, in the range of 10^9 bit/s for the retina, most human activities, from playing Tetris over memorization to typing, exhibit only about 10 bit/s of processing [ZM24]. The prevalence of parallel processing in sensory systems and serial processing in high-level cognition was identified as one of the major contributing factors. The significant fraction of time spent on reading in rapid memorization tasks, performed at up to 42 bit/s, illustrates that either there is more high-level parallelization than suggested or that, excluding perceptual time, the human brain can be way faster. If reading, associating and navigating were parallel, this would illustrate that even though humans seem to perform one action at the time, highest human processing speeds can be supported by an underlying parallelization of the involved brain regions. If the processes were instead sequential, the narrow difference between memorization and reading time would imply that the actual process of memorization reaches around 150 bit/s. Equally, any combination of the two options would explain part of the conundrum. It would be interesting to understand the speed of different cognitive subprocesses and how much parallelization exists during the use of memory palaces by top competitors.

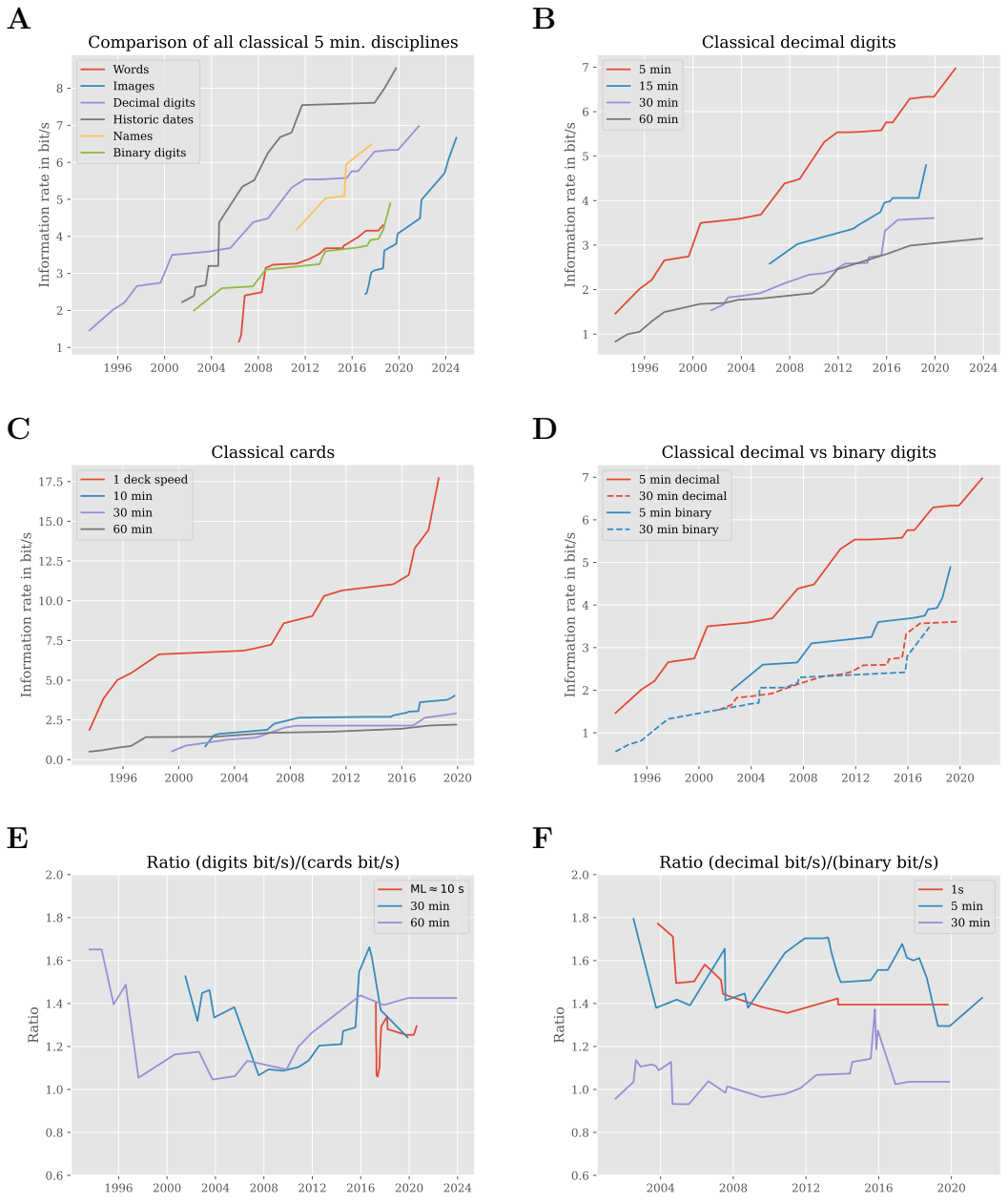

Figure 3: **A** Record development of all classical disciplines with 5 min memorization time. **B** Record development of classical disciplines based on memorizing long sequences of decimal digits. **C** Record development of classical disciplines based on playing-card memorization. **D** Record development of 5 and 30 min decimal- and binary-digits memorization. **E** The ratio of the information rates of all decimal-digits and cards disciplines of comparable length. Notably, this includes memorization times from 10 s to 60 min. The ML dataset started at a time when the disciplines were well-established, so the initial dip is not representative. **F** The ratio of the information rates of all decimal and binary digits disciplines of comparable length. Memorization times span from 1 s to 30 min.

It is plausible that the power-law decrease in the bit rate with time is a consequence of a larger time investment in revision and in better and more vivid associations to overcome the forgetting curve [MD15; AH11]. The probabilistic model suggests that the power law relates to the diminishing returns of the success probability of competitors to memorize certain information as time increases. During memorization, competitors may read the information only once at a slower pace, or revise twice or even several times in longer disciplines. As, additionally, the number of attempts on longer disciplines is smaller and we have no statistics on competitors errors, future controlled experiments will uncover the precise principles of the power law.

The consistent gaps in processing speed suggest that the difficulty of encoding systems for humans goes beyond the amount of conveyed information. The 52 options in playing cards add a complexity which is not present when memorizing decimal digits.

The 40% gap between binary and decimal digits in the 1 s and 5 min disciplines indicates that decimal digits strike a good balance between readability and complexity. We suppose that three factors explain the difference. First, the effect of word-length as more symbols are necessary to transfer the same information [Bar+14]. For instance, more saccades might be necessary to accumulate the information or more peripheral vision is used, which could explain lower information rates [SRJ11]. Second, the effect of reduced salience as visual binary information is more redundant than decimal information [Kow11]. Since binary saccade targets are less prominent, competitors might need to correct initial eye movements more often. Both effects can be experienced when reading the numbers

$$350286 = 1010101100001001110_2,$$

although the newest competition software tries to overcome the limitations through horizontal separation bars and colored cursor highlights [IAMc]. Third, competitors use systems for binary digits, which either carry less information per mnemonic image or were trained less. The most straightforward way to memorize binary digits is to convert three-digit binaries, say 010, into a single decimal digit, here 2. There are eight three-digit binary numbers, so that the strategy leads to a decrease of memorized information by

$$\frac{\log_2(8)}{\log_2(10)} \approx 0.9 \ .$$

Competitors start by learning a system to memorize decimal digits, as decimal information is more ubiquitous in real life and featured in other disciplines in speed-memory.com and classical competitions. Binary-specific systems, overcoming the above factor, might not be trained at all or not to the same level of proficiency.

The assimilation of processing speeds in the 30-min disciplines might be linked to the grading systems: decimal digits are presented in rows of 40 ($\log_2(10^{40}) \approx 133$ bits), whereas binary digits are given in rows of 30 ($\log_2(2^{30}) = 30$ bits). If the row is correct 40 respectively 30 points are awarded. If a single digit is different, points are halved and no points are awarded if more than two digits are incorrect. As practically all mnemonic systems for decimal and binary digits encode at least two digits with an mnemonic image [Wie25], an error results in the loss of all points for the row. Thus, the grading system punishes an error in decimal digits with a deduction of 133 bits and in binary digits only with 30 bits. In the half-an-hour discipline forgetting and errors might become more important than the ability to parse information. This might force competitors to revise or spend additional time in decimal compared to binary digits.

The information rates we calculated represent lower bounds on the speed of processing. In classical competitions records may contain errors, which lead to significant deductions in points. For associative disciplines such as names and faces or historic dates, we used a simplification that considers the associative elements to be as clearly distinguishable as possible. We did not model human memorization of written historic events, images or faces, which likely depends on manifold sentence/image properties [HE23] and domain expertise [NTO20; Gob98]. Equally, the use of memory palaces requires additional visuospatial and contextual information to be processed, which we did not consider. In general, the rules of memory competitions, such as grading schemes, are not precisely aimed at maximum information processing by competitors.

In conclusion, our systematic analysis reveals the following picture: over the shortest time scales, humans are able to memorize at up to 70 bit/s, but the rate decays with longer memorization times according to a power law. A major fraction of the required time in tasks of around 10 s can be explained by reading. This indicates that, if one excludes perceptual time, associations in memory might reach 100 bit/s. As the significant differences between encoding systems illustrate, humans are influenced by many aspects which go beyond the pure information-theoretic content of a tasks.

Acknowledgements I would like to thank Markus Meister, who drew my attention to the information rates of human performances.

This paper would not have been possible without the support of the memory sport community. I am grateful to Johannes Mallow and Simon Orton for sharing the dataset used to comment on past performances of top competitors during ML events. Johannes Mallow helped with several questions. Simon Orton kindly provided information on the generation of words and names in ML. I am thankful to Katie Kermode for discussions and for providing me with information on the generation of names and words in the classical competitions. Alex Mullen, Andrea Muzii, Don Michael Vickers and Enrico Marraffa generously gave insight into their performances.

Jan Rapp pointed out improvements of information calculations. Andreas Herz and Martin Stemmler contributed with stimulating discussions and helpful feedback on the manuscript.

Transparency Note The author has participated in classical and ML competitions using a three-digit number system and half two-card system.

4 Methods

4.1 Data acquisition

To analyze classical competitions, we used the world record histories [IAMd] from the statistics website [IAMb] of the International Association of Memory (IAM). The official online world records for ML can be found on [IAMa]. Results are only counted as world records if the result is obtained during an online competition with videos recorded from two different angles. Separate in-person records are kept but are outdated as in-person competitions have ceased to exist for ML since the COVID-19 pandemic. The world records for the 1 s and 4 s decimal and binary disciplines can be found on [Speb] and their development on [Spea].

Johannes Mallow (commentator on ML competitions [Mal] and former world champion) and Simon Orton (ML developer) provided us with a database of 191503 results of top competitors from the recent results on the ML website [ML] covering the period between June 2017 and January 2025. The dataset is used by commentators to assess the likelihood of participants to get a certain score during online competitions. We generated unofficial ML world records, including the ML cards performance of 11.89 s by Alex Mullen, by searching the database for the best results. A list of all used records can be found in Section 4.5.

4.2 Competitions and disciplines

A companion work elaborates on the used mnemonic techniques and provides additional background information on the competitions [Wie25].

4.2.1 Classical format

The classical format started with the first world championship in 1991. Over the years numerous annual championships were added. The competitions are held as a mental decathlon comprised of disciplines in Table 2.

There are three versions of the events taking place over the course of one, two or three days with correspondingly longer disciplines. Officially, the versions are called 'national', 'international' and 'world' standards. The three-day world standard is reserved for the world championships. The memorization times of the disciplines range from 5 to 60 minutes. For example, all standards feature 5 min decimal digits, and $15, 30$ and 60 min decimal digits respectively in the national, international and world format. The recall time is typically at least two or three times as long as the memorization time. To determine the overall winner, the raw scores of the disciplines are converted to a point scale with a system inspired by the track-and-field decathlon. More details can be found in the rulebook [IAM19] of the IAM.

4.2.2 Memory League

In 2015, Memory League was founded, first under the name Extreme Memory Tournament [ML]. There are fewer disciplines with much shorter memorization time, the focus is on one-vs.-

	memorization objective	memorization time [min]	grading scheme
decimal digits	a sequence of decimal digits	5, 15, 30, 60	rows of 40
binary digits	a sequence of binary digits	5, 30	rows of 30
auditory digits	a sequence of decimal digits, read aloud at one digit/s	up to 8	counts until first error
cards	a sequence of playing cards in the form of successive scrambled decks of 52 playing cards	10, 30, 60	rows (decks)
speed cards	the sequence of one scrambled deck of 52 playing cards as quickly as possible	at most 5	essentially perfect
words	a sequence of words	5, 15	rows of 20
images	a sequence of images with arbitrary motives in various art-styles	5	rows of 5
historic dates	associate years between 1000 and 2100 to fictional events	5	deduction (guessing not allowed)
names and faces	associate international names to faces (first + last name)	5, 15	no deduction (guessing allowed)

Table 2: Overview of the disciplines in classical competitions. The first seven disciplines involve the memorization of long sequences. Decimal and binary digits, cards and words are presented in rows of k items: a single error in a row leads to the row being counted as $k/2$, whereas with two or more errors no points are being awarded. Speed cards is the only discipline in which the time rather then the amount counts. The number of cards counts until the first error - however only an error-free performance leads to a competitive score. In images recall competitors need to indicate the previous permutation of the five images in a row with a deduction of one point, if an error is made. As images rows contain only five items, strategies which do not involve knowing the overall sequence, but only the sequence per row, are possible. Historic dates and names and faces are associative disciplines. In recall, the events or the faces are given in a random permutation and competitors need to reconstruct the year respectively the name. The disciplines differ in our ability to estimate the information rate processed by humans in bit/s. For the purely number or card based disciplines entropies are straightforward to calculate. For images and historic dates we use theoretic lower-bounds, whereas for word and names we employ estimates based on database samples (Section 4.3).

one matches and top competitors optimize the time for a certain number of items, rather than optimizing the number of items per time.

ML features the six disciplines decimal digits, cards, words, national names, international names, images (Table 3). The memorization time lasts up to one minute, recall up to four minutes and participants are able to end both prematurely. Tournaments are based on matches in which two players compete against each other. To evaluate the winner of a given discipline, the number of items memorized/accuracy counts. If the accuracy of both competitors is equal, the player wins who ended the memorization period faster. Except in the international names discipline, the best participants typically aim to memorize all information. The grading system then amounts to: whoever made fewer errors wins and, if both have made the same number of errors, the person with the faster memorization time wins. Competitors can only see the time and score of the opponent after recall, allowing for different strategies. For example, competitors may attempt a fast perfect score while risking errors, or they can comfortably memorize everything in one minute, and hope for an error of the opponent.

4.2.3 Speed-memory.com

The speed-memory.com competitions consist of six disciplines: 1 s and 4 s decimal and binary digits and the matrices and colored shape disciplines. The website [Speb] is in Spanish and the

	memorization objective
words	a sequence of 50 words
images	a sequence of 30 real-life photos with arbitrary motives
numbers	an 80 digit decimal number
cards	the sequence of a scrambled deck of 52 playing cards
national names	associate 30 faces to national names (only first name)
international names	associate 30 faces to international names (only first name)

Table 3: Overview of the disciplines in Memory League (ML). Compared to classical competitions the time scale is shorter. The memorization time may last at most one minute compared to at least five minutes in classical competitions with the exception of speed cards. Instead of optimizing the amount of information in a given time, competitors try to memorize a given set of information as quickly as possible.

competitive scene centered around Ramón Campayo and his students. At the time this article was written, it was unclear from the website whether there still was a competitive scene as there were no upcoming events and several sections of the website seemed outdated. In the 1 s and 4 s disciplines competitors see all information at once, but are allowed to place the decimal and binary digits in particular groups arranged across the screen.

4.3 Information content of the different disciplines

Shannon's entropy is defined as

$$H(\mathbb{P}) = -\sum_{x \in \Omega} \log_2(\mathbb{P}(x))\mathbb{P}(x)$$

for a probability distribution $\mathbb{P}$ on a discrete underlying event space Ω. Marking rules and the high number of possibilities prevent guessing of competitors. Therefore, if we calculate the entropy associated to the recalled items, it will correspond to transmitted information. Ideally, we would prefer to have a distribution of answers of a competitor. The records present only one realization, but at least the disciplines consist out of a sequence of repetitive microtasks such as individual digits. For all disciplines, our calculations are based on the idea to provide a lower bound on the information necessary to explain the human performance.

For example, in classical digits, cards, words and images disciplines we assume performances are perfect, even though the row grading scheme leads to raw scores which are lower than the number of correctly recalled items. The disciplines are given in rows of k items. A single error in a row leads to the row being counted as $k/2$, whereas with two or more errors no points are being awarded. If the last row is only recalled up to a certain item, indicating that the competitor has not memorized further, the rule is applied to the partial row.

For words and names, we use estimates based on large samples of databases. For the associative aspect in names, images and historic dates, we use simplified lower bounds, as any other derivation would require additional assumptions on human cognition. We do not incorporate any of the underlying mental processes, such as the memory palace, in our calculations. We denote entropy associated to achieving a raw score n in a discipline by $H(\text{discipline}, n)$.

4.3.1 Names and faces

During memorization in classical competitions, competitors are given faces with a first and a last name. In the recall period, faces are scrambled and competitors need to recall the names. A point is awarded for every correctly retrieved name, so two points are possible per face. This implies that a raw score of n corresponds to memorizing at least $n/2$ faces. In ML national and international names, 30 faces are given with only a first name.

IAM names are generated as follows:

- Names are uniformly chosen from eight different 'regions', which cover different language branches.
- In 5 min names, there are 12 "long" names of eight letters or more and in the 15 min version 24 long names. The long names are chosen uniformly among first and last names.

In the memorization phase the number of names equals the world record plus 20%. As the current 5 min world record is 105, we assume a fraction 12/126 of long names. Katie Kermode, who co-developed the latest IAM training software [IAMc], provided us with a list of the numbers of short/long, male/female per region. This allowed us to calculate the entropy carried by IAM names as

$$H(\text{First name}) = 13.34\ ,$$
$$H(\text{Second name}) = 12.24\ .$$

A score of n requires at least $n/2$ faces to be distinguished in recall. Assuming that each of the faces was instead a number, or as clearly distinguishable as possible, one can see that the lower bound for the information of this association task is $\log_2(n/2)$. During record performances almost all of the given information is memorized, allowing us to assume a similar amount of first and last names. We obtain the overall lower estimate

$$H(\text{Classical names}, n) \approx n\Big(\frac{H(\text{First name}) + H(\text{Second name})}{2} + \log_2(n/2)\Big).$$

ML does not share the precise names distribution, but Simon Orton, the developer of ML, kindly applied the entropy formula for us. Both male and female names have an entropy of

$$H(\text{ML national name}) = 8.8\ .$$

This allows us to calculate the entropy associated to a national names performance in ML as

$$H(\text{ML national names}, n) \geq n\big(H(\text{ML national name}) + \log_2(n)\big)$$

for $0 \leq n \leq 30$. International names consists out of all 41 national names databases combined. Simon Orton provided us again with the result of the entropy formula

$$H(\text{ML international name}) = 14.7\ .$$

Similarly to national names,

$$H(\text{ML international names}, n) \geq n\big(H(\text{ML international name}) + \log_2(n)\big)$$

for $0 \leq n \leq 30$.

4.3.2 Words

Words is the only discipline for which no database for IAM competitions exists. As competitors memorize words in their native language, it is easier for organizers to provide translations for the then limited task set. As such, performances are slightly harder to compare than in other disciplines. Nevertheless, the discipline can be trained on the IAM training website [IAMc]. We assume that players would have complained if the training mode was too easy and used the list logged in the console of the browser. It contains 1683 concrete nouns, 829 abstract nouns and 763 verbs with some words classified for several categories. The composition of the sequence of words in competitions is 80% concrete nouns, 10% abstract nouns and 10% verbs [IAM19]. Accounting for words being in several categories, this leads to the estimate

$$H(\text{Classical word}) = 8.9\ .$$

We then use the formula

$$H(\text{Classical words}, n) \approx nH(\text{Classical word}).$$

Simon Orton informed us that ML words are uniformly chosen among the ML words database. We found more than 3900 different words, giving us the lower bound

$$H(\text{ML word}) \geq \log_2(3900) \approx 11.93\ .$$

This allows us to calculate the entropy associated to a words performances in ML as

$$H(\text{ML words}, n) \geq nH(\text{ML word}) = n \cdot 11.93\ .$$

for $0 \leq n \leq 50$, as 50 words are given.

4.3.3 Images

In this discipline, the task is to recreate the original order of a sequence of images. While the scores reach several hundred in the classical format, it is important to note by what system the images discipline is marked. Images are presented in rows of five, and in recall each row of five images is shown in a random order and participants need to indicate the previous order by numbers. There is a deduction of one point per incorrect row preventing guessing. Enrico Maraffa, the world record holder in IAM images, confirmed in personal communication that he is not using memory palaces, but only links. Therefore, he would be able to recall the same information if the rows were scrambled. As he is able to identify the rows, similar to the associative aspect in names, we add $\log_2\left(\frac{n}{5}\right)$ for the row identification to arrive at

$$H(\text{Classical images (IAM)}, n) \geq \frac{n}{5}\Big(\log_2(5!) + \log_2\left(\frac{n}{5}\right)\Big).$$

In ML, participants see 30 images and need to reassemble them during recall. Thus, the entropy of the ML images discipline is

$$H(\text{ML images}, n) \geq \log_2(n!)$$

for $0 \leq n \leq 30$.

4.3.4 Digits

The calculation of the number of bits conveyed by decimal digits and binary digits is straightforward:

$$\begin{aligned} H(\text{Binary digits}, n) &= n, \\ H(\text{Decimal digits}, n) &= n\log_2(10), \\ H(\text{ML digits}, n) &= n\log_2(10), \qquad \forall 0 \leq n \leq 80. \end{aligned}$$

4.3.5 Historic dates

The objective is to memorize the year between 1000 and 2100 of fictional historic dates. In recall, the fictional events are given in a random order without years. The entropy carried by the year of a single memorized historic date is given by $\log_2(1100) = 10.1$. We again need to account for the associative component by adding $\log_2(n)$. In reality encoding will be less efficient. We can thus provide a lower bound to the information processed in the discipline by

$$H(\text{Historic dates}, n) \geq n\big(\log_2(1100) + \log_2(n)\big).$$

4.3.6 Cards

The entropy for a only a subset of the 52 playing cards being correct is given by

$$H(\text{One deck, } n \text{ out of 52 correct}) = H(\text{ML Cards}, n) = \log_2\left(\frac{52!}{(52-n)!}\right).$$

The discipline is subject to the row grading scheme with the decks serving as rows. If the competitor achieved a raw score n, we assume that $\lfloor \frac{n}{52} \rfloor$ decks were memorized correctly and that any remainder $n \mod 52$ originates from the memorization of part of a deck at the end of the recalled sequence

$$H(\text{Classical cards}, n) = \left\lfloor \frac{n}{52} \right\rfloor \log_2(52!) + \log_2\left(\frac{52!}{(52 - n \mod 52)!}\right).$$

4.4 Probabilistic model

We present a short probabilistic model relating the information-rate power law to the probability of perfect memorization in short tasks. The main idea is that competitors try to remain error-free and that a long discipline can be decomposed into shorter, independent tasks. It necessarily neglects many of the aspects described in the previous sections. We model one of the sequential disciplines. Suppose that competitors try to achieve perfect memorization to avoid incurring any of the rows being awarded no points. To do so, they aim to achieve perfect memorization with a certain probability

$$\mathbb{P}(\text{all correct}) \approx p \in (0,1).$$

We subdivide the sequence into groups of k items, each with entropy $H(\text{item})$. Let $A(k,t)$ be the event that a competitor attempts and correctly recalls k elements in t seconds memorization time excluding the reading time r for the group. Let T be the overall time of the discipline in seconds. Assuming that groups of k elements are independent of each other, we could then write

$$p \approx \mathbb{P}(\text{all correct}) = \mathbb{P}\big(A(k,t)\big)^{T/(t+r)} \Rightarrow \mathbb{P}\big(A(k,t)\big) \approx p^{(t+r)/T}.$$

The official records suggest that the information rate scales as a power law of the discipline, so we obtain

$$aT^b = R(T) = \frac{kH(\text{item}) \cdot \frac{T}{(t+r)}}{T} = \frac{kH(\text{item})}{(t+r)} \Rightarrow T = \Big(\frac{kH(\text{item})}{a(t+r)}\Big)^{1/b}.$$

This implies that

$$\mathbb{P}\big(A(k,t)\big) = p^{(t+r)\left(\frac{a(t+r)}{kH(\text{item})}\right)^{1/b}}.$$

The prediction $\mathbb{P}\big(A(k,t)\big)$ for fixed k can be found in Figure 2. If, alternatively, one fixes t, the model predicts that the probability to memorize all k items resembles a reversed logistic growth curve. For a small number k of items the probability is close to one. As the number of items k increases, the decay of the probability is first slow, accelerates and then decelerates, so that the probability asymptotically converges to zero.

4.5 List of records

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical names	Memorization time 15 min	224 names	4.88	Katie Kermode
	Memorization time 5 min	105 names	6.48	Katie Kermode
ML international names (official)	30 names	58.51 s	10.05	Katie Kermode
ML international (unofficial)	30 names	54.1 s	11.08	Katie Kermode
ML national names (official)	30 names	27.91 s	14.73	Matteo Cillo
ML national names (unofficial)	30 names	20.96 s	19.62	Matteo Cillo

Table 4: Names world records.

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical words	Memorization time 15 min	318 words	3.15	Katie Kermode
	Memorization time 5 min	145 words	4.3	Yanjaa Wintersoul
ML words (official)	50 words	32.41 s	18.4	Don Michael Vickers
ML words (unofficial)	50 words	27.95 s	21.34	Don Michael Vickers
	Youtube training [Vic]	12.77 s 45/50	42.04	Don Michael Vickers

Table 5: Words world records.

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical images	Memorization time 5 min	711	6.66	Enrico Marraffa
ML images (official)	30 images	8.6 s	12.52	Matteo Cillo
ML images (unofficial)	30 images	7.7 s	13.99	Matteo Cillo

Table 6: Images world records.

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical digits	Memorization time 60 min	3412	3.15	Orkhan Ibadov
	Memorization time 30 min	1955	3.61	Sylvain Arvidieu
	Memorization time 15 min	1300	4.8	Munkshur Narmandakh
	Memorization time 5 min	630	6.98	Andrea Muzii
ML digits (official)	80 decimal digits	11.45 s	23.2	Alex Mullen
ML digits (unofficial)	80 decimal digits	10.82 s	24.56	Andrea Muzii
	Youtube training [Muz]	9.75 s 80/80	27.26	Andrea Muzii
Speed Memory	Memorization time 4 s	31	25.75	Joaquin Garcia
	Memorization time 1 s	21	69.76	Ramón Campayo

Table 7: Decimal digits world records.

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical binary	Memorization time 30 min	6270	3.48	Munkshur Narmandakh
	Memorization time 5 min	1467	4.89	Munkshur Narmandakh
Speed Memory	Memorization time 4 s	100	25	Joaquín Garcia
	Memorization time 1 s	50	50	Joaquín Garcia

Table 8: Binary digits world records.

	Description	WR	$[\frac{\text{bit}}{s}]$	Record holder
Classical cards	Memorization time 60 min	1829	2.21	Orkhan Ibadov
	Memorization time 30 min	1202	2.9	Sylvain Arvidieu
	Memorization time 10 min	550	4.02	Andrea Muzii
	One deck, 52 playing cards	12.74 s	17.71	Shijir-Erdene Bat-Enkh
ML cards (official)	One deck, 52 playing cards	12.25 s	18.41	Alex Mullen
ML cards (unofficial)	One deck, 52 playing cards	11.89 s	18.97	Alex Mullen

Table 9: Cards world records.

References

[AH11] L. Averell and A. Heathcote. "The form of the forgetting curve and the fate of memories". *Journal of Mathematical Psychology* 55.1 (2011). Special Issue on Hierarchical Bayesian Models, 25–35.

[Arta] Art of Memory Forum. *Andrea Muzii three-digit/PAO cards*. URL: `https://forum.artofmemory.com/t/considerations-about-the-new-world-memory-champion-2019-andrea-muzii/44661` (visited on 02/26/2025).

[Artb] Art of Memory Forum. *History of the major system*. URL: `https://artofmemory.com/blog/history-of-the-major-system/` (visited on 01/23/2025).

[Bar+14] J. Barton et al. "The word-length effect in reading: A review". *Cognitive Neuropsychology* 31.5-6 (2014), 378–412.

[Gob98] F. Gobet. "Expert memory: a comparison of four theories". *Cognition* 66.2 (1998), 115–152.

[Gri22] J. Grimwade. "Memory in Roman Oratory: Theory and Practice". PhD thesis. Apollo - University of Cambridge Repository, 2022.

[Hau+12] O. Hauk et al. "The time-course of single-word reading: Evidence from fast behavioral and brain responses". *NeuroImage* 60.2 (2012), 1462–1477.

[HE23] T. Hagen and T. Espeseth. "Image memorability prediction with vision transformers". *arXiv preprint arXiv:2301.08647* (2023).

[IAMa] IAM. *ML records*. URL: `https://iam-stats.org/records_ml_online.php` (visited on 01/17/2025).

[IAMb] IAM. *Statistics*. URL: `https://iam-stats.org` (visited on 10/20/2024).

[IAMc] IAM. *Training website*. URL: `https://www.iamwmc.com/competition/training.html` (visited on 01/13/2025).

[IAMd] IAM. *World records*. URL: `https://iam-stats.org/records.php` (visited on 01/17/2025).

[IAM19] IAM. *Rulebook*. 2019. URL: `https://www.iam-memory.org/wp-content/uploads/2019/11/IAM_Rulebook_Nov2019.pdf/` (visited on 09/27/2024).

[Kow11] E. Kowler. "Eye movements: The past 25 years". *Vision Research* 51.13 (2011). Vision Research 50th Anniversary Issue: Part 2, 1457–1483.

[Mal] J. Mallow. *MemorysportTV*. URL: `https://www.twitch.tv/memorysportstv` (visited on 02/27/2025).

[MD15] J. M. J. Murre and J. Dros. "Replication and Analysis of Ebbinghaus' Forgetting Curve". *PLOS ONE* 10.7 (July 2015), 1–23.

[ML] ML. *Memory League*. URL: `https://memoryleague.com/#!/home` (visited on 01/17/2025).

[Muz] A. Muzii. *80/80 decimal digits in 9.75s*. URL: https://www.youtube.com/shorts/bLe0cUOolUw (visited on 01/31/2025).

[NTO20] D. G. Nagy, B. Török, and G. Orbán. "Optimal forgetting: Semantic compression of episodic memories". *PLoS Computational Biology* 16.10 (2020), e1008367.

[PK57] J. R. Pierce and J. E. Karlin. "Reading rates and the information rate of a human channel". *The Bell System Technical Journal* 36.2 (1957), 496–516.

[PS18] R. W. Proctor and D. W. Schneider. "Hick's law for choice reaction time: A review". *Quarterly Journal of Experimental Psychology* 71.6 (2018), 1281–1299.

[Spea] *Speed memory record development*. URL: https://www.recordholders.org/en/records/quick-memory/quick-memory.html (visited on 02/28/2025).

[Speb] *Speed memory records*. URL: https://www.speed-memory.com/records/ (visited on 02/28/2025).

[SRJ11] H. Strasburger, I. Rentschler, and M. Jüttner. "Peripheral vision and pattern recognition: A review". *Journal of Vision* 11.5 (2011), 13–13.

[Tho16] H. Thomson. *The man who thinks like sherlock holmes*. 2016. URL: https://www.bbc.com/future/article/20160411-the-man-who-thinks-like-sherlock-holmes (visited on 03/18/2025).

[Vic] D. M. Vickers. *45/50 words in 12.77s*. URL: https://www.youtube.com/watch?v=Vkgo89HH3C4 (visited on 01/31/2025).

[WE91] J. T. Wixted and E. B. Ebbesen. "On the Form of Forgetting". *Psychological Science* 2.6 (1991), 409–415.

[Wie25] B. Wiederhold. "An introduction to memory competitions, records and techniques" (2025). In preparation.

[ZM24] J. Zheng and M. Meister. "The unbearable slowness of being: Why do we live at 10 bits/s?" *Neuron* 113.2 (2024), 192–204.