nature human behaviour

ARTICLES
https://doi.org/10.1038/s41562-018-0474-5

# The universal decay of collective memory and attention

**Cristian Candia[1,2,3]*, C. Jara-Figueroa[1], Carlos Rodriguez-Sickert[3], Albert-László Barabási[2] and César A. Hidalgo[1]***

**Collective memory and attention are sustained by two channels: oral communication (communicative memory) and the physical recording of information (cultural memory). Here, we use data on the citation of academic articles and patents, and on the online attention received by songs, movies and biographies, to describe the temporal decay of the attention received by cultural products. We show that, once we isolate the temporal dimension of the decay, the attention received by cultural products decays following a universal biexponential function. We explain this universality by proposing a mathematical model based on communicative and cultural memory, which fits the data better than previously proposed log-normal and exponential models. Our results reveal that biographies remain in our communicative memory the longest (20–30 years) and music the shortest (about 5.6 years). These findings show that the average attention received by cultural products decays following a universal biexponential function.**

In what is probably Pablo Neruda's most famous poem—'Poema 20'—he wrote: 'Es tan corto el amor, y tan largo el olvido' (Love is so short, forgetting is so long). Neruda's words express elegantly the fact that, when people are in love, they are constantly thinking of their loved ones, but once love fades, memories fade too. Inspired by Neruda, we ask whether society also experiences the two phases of memory: an initial phase of high attention, followed by a longer and slower phase of forgetting. In fact, there is a vast literature suggesting that this might be the case, as collective memory is acknowledged to be a combination of two distinct processes[1–11]: communicative memory, normally sustained by the oral transmission of information, and cultural memory, which is sustained by the physical recording of information. This literature can provide inspiration for the construction of generative models for the attention received by cultural products.

Despite this progress, the theory of collective memory and attention is short on quantitative models that would allow us to connect it empirically to large-scale data, such as the data developed in the literature of knowledge diffusion. Indeed, the literature on knowledge diffusion models the adoption and diffusion of cultural content (Fig. 1a) as a combination of two processes[12–18]: preferential attachment (Fig. 1b) and temporal decay (Fig. 1c). Preferential attachment[19,20], or cumulative advantage[21–23], refers to a process in which attention begets attention. Think of two scientific papers, one with 10,000 citations and another one with 100. The probability that the first paper receives a new citation is larger than the second one, simply because more people already know about it. This preferential attachment process needs to be properly discounted to measure temporal decay.

Recently, models combining preferential attachment and temporal decay have described the decay of attention (mostly paper and patent citations) using exponential and log-normal functions[12,13]. These models agree on the idea that attention should be modelled using a combination of preferential attachment (Fig. 1b) and time decay (Fig. 1c). Yet, there is no consensus about the shape of the decay function or its universality across various cultural domains.

Here, we use data on scientific publications, patents, songs, movies and biographies to test the hypothesis that the decay of the attention received by these cultural products involves the decay of both communicative and cultural memory. Owing to the properties ascribed to each type of memory—communicative memory being short lived compared to cultural memory[24]—we expect that the attention received by collective memory should decay fast at first, whereas that of cultural memory should follow a softer and longer decline. We formalize these ideas by constructing a mathematical model that predicts a biexponential decay function and validate it by showing that it is statistically better at explaining the empirically observed decay of attention than the exponential[13] and log-normal[12] functions used in the previous literature. This finding validates our hypothesis that the decay of the attention received by human collective memory is a process that results from the interplay between both communicative and cultural memory. The model also allows us to separate both mechanisms and generalizes well to multiple data sets, suggesting that it captures a universal feature of the decay of human collective memory.

## Collective memory and attention

Collective memories are sustained by communities, which could be as large as all of the speakers of a language or as small as a family. During the past century, scholars studying collective memory have advanced a large number of definitions, models and processes, helping to characterize different forms of collective memory and the mechanisms that contribute to their preservation[25].

Psychologists have explored both top-down and bottom-up approaches to memory formation and retention. Top-down approaches focus on how familiarity[26,27], narrative templates[5,28] and cultural attractors[29–31] contribute to the retention and formation of collective memories. Familiarity increases the memorability of events, even causing false memories, such as that of people

[1]Collective Learning Group, The MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Network Science Institute, Northeastern University, Boston, MA, USA. [3]Centro de Investigación en Complejidad Social (CICS), Facultad de Gobierno, Universidad del Desarrollo, Santiago, Chile. *e-mail: ccandiav@mit.edu; hidalgo@mit.edu

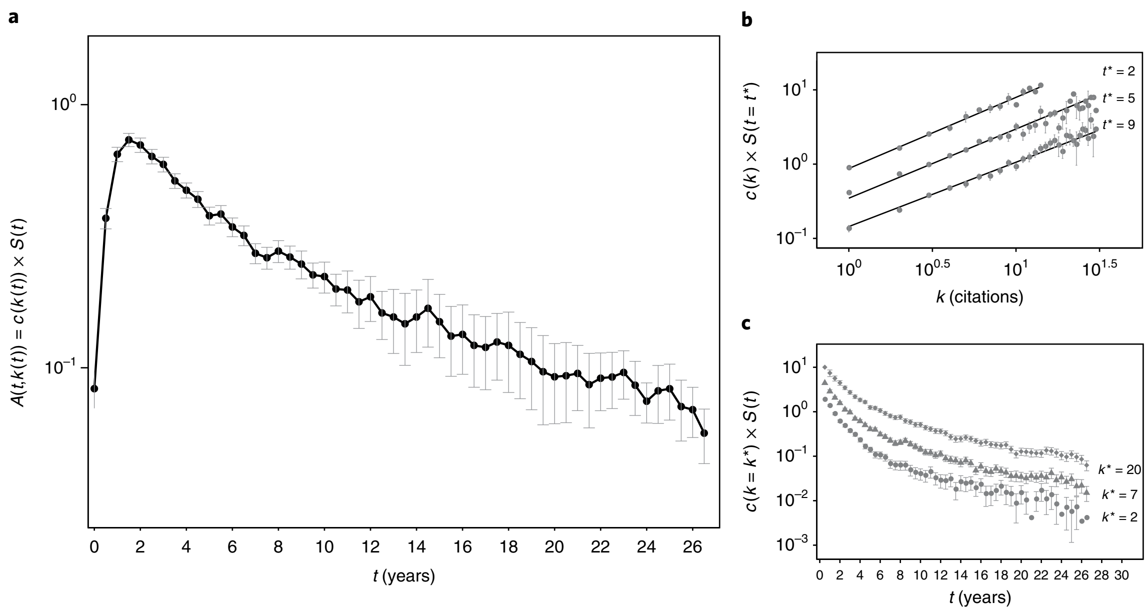


**Fig. 1 | Universal patterns in the decay of human collective memory. a**, Average number of citations received each semester by papers published in *Physical Review B* ($A(t)$). **b**, Average number of citations received by a paper as a function of the cumulative citations received by that paper ($\propto c(k)$). Different curves represent different ages. **c**, Average number of citations received by papers with the same number of cumulative citations as a function of their age ($\propto S(t)$). Different curves represent groups of papers with a different number of total citations. Error bars represent standard errors.

identifying Alexander Hamilton as a US president[26]. Narrative templates, which are schemata that people use to describe multiple historical events, can also shape memories, such as the memory of Russian exceptionalism that emerges from the narrative template of invasion, near defeat and heroic triumph[5]. Cultural attractors, such as repetitive children songs or count-out rhymes, can increase the preservation of memories across generations[27].

Bottom-up approaches focus on how micro-level psychological processes can shape social outcomes[25]. For instance, forgetting can be induced through the selective retrieval of events, an effect known as retrieval-induced forgetting[32–34]. In addition, social affinities, such as belonging to the same social group, can increase the mnemonic power of conversations[35–38]. For instance, people are motivated to create coalitions[39] and shared realities with those that they perceive as belonging to their own group[36].

Scholars in computational social sciences have followed a different approach, focusing on how collective memory is expressed and created in the consumption of cultural content, from Wikipedia page views[11,40–43] to paper and patent citations[12,13,44,45]. Of course, these online and offline metrics are not direct measures of collective memory or attention, they are measures of the spillovers of attention that result in online searches or references. The idea is that movies, songs or papers that are being talked about are of heightened interest, and hence, lead people to consult various data sources. When these cultural products move away from communicative memory, they lose the intense attention that they had when they were being talked about.

Unfortunately, these aggregate approaches cannot distinguish between different forms of memory or attention loss, such as interference, suppression or inhibition. They only provide an aggregate picture of the attention lost through all of these channels.

Nevertheless, the computational social science approach is closer to the definition of collective memory given by Jan Assmann[2,4], which focuses on the cultural products that communities or groups of people remember. Assmann—even though he focused on long-lived inter-generational memories—distinguishes between modes of potentiality and actuality: potentiality being the existence of a record (an old book in a library's basement), and actuality being the attention received by that record when it becomes relevant to the community. The computational social science literature has focused on the use of big data to study the actuality of memories and the effects of language, technology, accomplishments and triggers in the dynamics of collective memory and attention. For instance, historical figures born in countries with languages that are often translated to other languages receive more online attention than comparable historical figures born in less-frequently translated languages[46]. Changes in communication technologies, such as the rise of the printing press, radio and television, have also been shown to affect attention as they correlate with changes in the occupations of the people entering biographical records[41]. The edits and attention received by events in Wikipedia have also been seen to increase with related exogenous events[11,43], such as natural and human-made disasters, accidents, terrorism and during anniversaries or commemorative events[47]. Moreover, the online attention received by past sports figures—a measure of their prevalence in present-day memory—has been shown to correlate with an age-discounted measure of performance[40,48], meaning that memorability and attention—at least in athletic activities—correlate with merit.

The approach presented in this paper is related more closely to the computational social science strand of literature, as it uses cultural consumption data to study the dynamics of the attention received by the previously described cultural products and biographies. Yet, it is also an approach that is not completely unrelated to the psychological strand. By studying the dynamics of consumption of these cultural products, from songs to scientific papers, we are exploring a form of selective retrieval, albeit not focused on how this selective retrieval shapes collective identity, but on its average temporal dynamics. Moreover, by proposing a model that describes the dynamics of attention, we are undertaking a bottom-up approach to the modelling of collective memory and attention. Finally, by looking at multiple cultural domains, we can explore the universality of average decay functions, rather than focusing on the mechanisms that make some events more or less memorable.

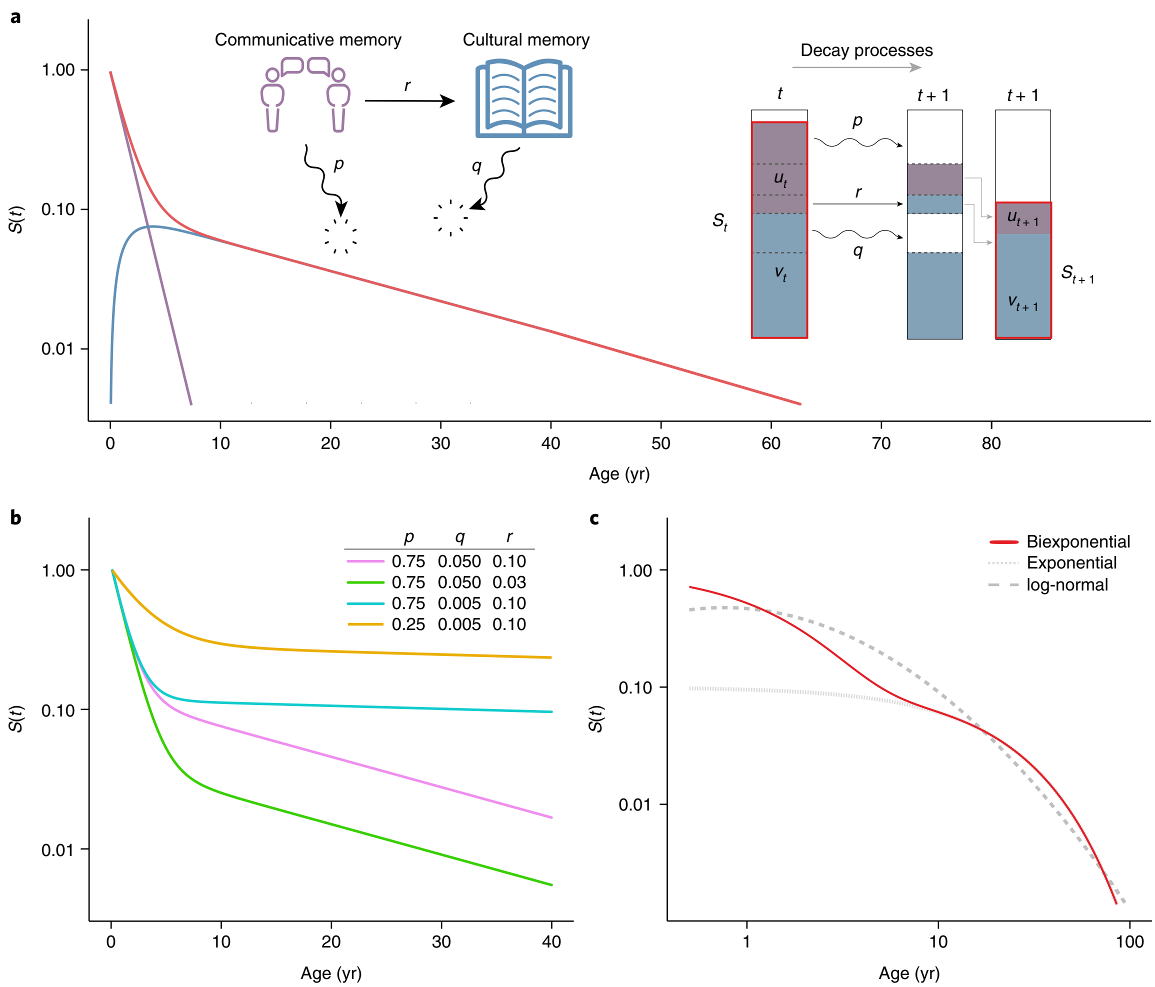


**Fig. 2 | Scheme of the collective memory model. a**, The $y$ axis represents the normalized current level of attention received by a group of comparable cultural pieces. The $x$ axis represents the age of the cultural pieces. The red curve shows the biexponential function predicted by our model in log-lin scale. The blue and purple curves show the two exponentials of communicative and cultural memory, respectively. The inset illustrates the basic mechanics of the model. At any time point $t$, the total memory is the sum of communicative memory $u$ and cultural memory $v$. Both communicative and cultural memory decay with their own respective decay rates $p+r$ and $q$, and cultural memory grows with $r$. **b**, The biexponential model (equation (6)) for various parameters $p$, $q$ and $r$ can account for a wide range of decays. **c**, Comparison between the biexponential model and the exponential and log-normal models in log-log scale.

## Results

The literature on collective memory[1,24,49] suggests that the decay of the attention received by a cultural product involves two mechanisms: an initial fast decay—a signature of communicative memory—followed by a softer decline—resulting from cultural memory (Fig. 2a). Using the distinction between communicative and cultural memory[3,4,24,49], we propose a model in which cultural memory and communicative memory co-exist, but decay at different rates. The decay of both types of memory, especially cultural memory, should be understood in relative terms: the share of the current attention occupied by a cultural product may stay the same, but because the total memory is growing, more products are created each time. Hence, the relative share of the current attention ($S(t)$) assigned to the said product will decay.

We model the attention received by a cultural product using several simplifying assumptions. First, we assume that the current attention, $S(t)$, of a cultural product is the sum of the attention it garners from both communicative memory $u$ and cultural memory $v$. Hence, at any given time $S(t)=u(t)+v(t)$ (Fig. 2a). Second, we assume that communicative and cultural memory decay, in relative terms, independently with decay rates $p+r$ for communicative memory and $q$ for cultural memory, and that information flows from communicative memory to cultural memory at a rate $r$. Many processes are captured by the parameters $p$, $r$ and $q$, perhaps the most straightforward one is that because the total size of cultural memory is growing, the relative share occupied by a certain cultural product will shrink, which is captured in $p$. We acknowledge that these assumptions cannot capture the full complexity of the processes by which communicative and cultural memory decay, nor their interactions. The communicative and cultural memory may feed on each other in more-complex ways than the assumed linear form ($r$). We adopt these simplifying assumptions with the goal of providing a tractable model with as few parameters as possible that can be used to capture the leading forces that govern the dynamics of attention received by a cultural product. Given these assumptions, communicative memory decays as $u(t+1)=(1-p)u(t)-ru(t)$ and cultural memory as $v(t+1)=(1-q)v(t)+ru(t)$, together defining the following system of differential equations (see the derivation of the model in the Methods section (under 'Model'):

$$S(t)=u(t)+v(t) \tag{1}$$

$$\frac{du}{dt} = -(p+r)u \tag{2}$$

$$\frac{dv}{dt} = -qv + ru \tag{3}$$

We set the initial communicative memory $u(t=0)=N$ and we assume that, at the beginning of the process, there is no cultural memory associated to a new cultural product ($v(t=0)=0$), although there are alternatives ways to initialize the model that does not change its aggregate behaviour (Supplementary Model).

Using the initial conditions, we find that the solution of the equation system (equations (1)–(3) is the biexponential function:

$$u(t) = Ne^{-(p+r)t} \tag{4}$$

$$v(t) = \frac{Nr}{p+r-q}(e^{-qt} - e^{-(p+r)t}) \tag{5}$$

$$S(t) = \frac{N}{p+r-q}[(p-q)e^{-(p+r)t} + re^{-qt}] \tag{6}$$

Figure 2b illustrates $S(t)$ for different values of the parameters, with $N=1$, and Fig. 2c compares the biexponential function with the exponential[13,14] and log-normal[12] decay functions explored previously in the literature (Supplementary Notes 1–3).

We bring the biexponential model to our data by comparing it with the decay functions observed for paper and patent citations, and for the current online attention of past songs, movies and biographies, with a comparable level of accomplishment (Table 1). In the case of papers and patents, we group papers and patents with a similar number of cumulative citations. In the case of songs, movies and biographies, these comparable sets are built into our selection criterion, as we study only songs that reached the Billboard ranking, biographies of award-winning athletes and movies that have received over 1,000 votes on the Internet Movie Database (IMDB). By respectively grouping papers, patents, songs, movies and biographies, with a similar level of accomplishment, we control for differences in preferential attachment, allowing us to isolate the temporal decay of collective memory statistically (see 'Data' in the Methods section and Supplementary Methods).

Figure 3 shows the average number of new citations obtained by scientific papers (A, B, C and D) and patents (E and F) for different levels of accumulated citations $k$. The red lines show the fit of the biexponential model, whereas the dashed and dotted lines capture, respectively, the log-normal and exponential decays used in refs [1213]. In all cases, we find, after choosing papers and patents with the same level of cumulative citations, positive differences of the corrected Akaike's information criterion (AICc) between the log-normal and biexponential models (Fig. 4a) and positive differences of $R^2$ measures between the log-normal and biexponential models (Fig. 4c). This suggests that the biexponential model captures the temporal pattern of human collective forgetting accurately (see 'Goodness of fit' in the Methods and Supplementary Tables 1–3 for a comparison of the data on all years, journals and categories). More importantly, in several of these empirical curves, the shoulder of the biexponential curve is clearly visible, allowing the model to help to unveil the point at which cultural memory takes over communicative memory.

We observe a similar behaviour when we apply the biexponential model to data on music, movies and biographies. As we lack time-series data for these three sources, we look at the present-day online attention to music (Fig. 3g), movies (Fig. 3i) and biographies (Fig. 3j) as a function of their age. For songs, we determine age using the year they first reached the Billboard ranking. For movies, we calculate age using their release year. For the biographies of athletes, we use as the age of the accomplishment the time when they were introduced in their respective international rankings. Once again, when we compare our model with the previously proposed log-normal and exponential models, we find that the biexponential model provides a more accurate fit to the data, owing to its ability to capture the initial fast decay of communicative memory together with the slow decay of cultural memory. Furthermore, it visibly captures the transition from communicative to cultural memory.

**Table 1 | Cultural products and their measurements of present-day levels of attention (current attention) and measurements to account by cumulated advantage effect (accomplishment)**

| Cultural products | Attention metric | Preferential attachment metric |
|---|---|---|
| APS papers | Citations received in the past six months | Cumulative citations |
| USPTO patents | Citations received in the past six months | Cumulative citations |
| Music | Spotify popularity and Last.fm play counts | Entered at least once in the Hot-100 Billboard ranking |
| Movies | Trailer play counts in YouTube | More than 1,000 votes on IMDB |
| Biographies | Wikipedia page views | Highly performing athletes in tennis, basketball and the Olympics |

Together, the data on papers, patents, songs, movies and biographies show that this biexponential decay is universal across all domains. Yet, the parameters of the decay function are different for papers, patents, songs, movies and biographies. Thus, we compared the model parameters ($p$, $q$, $r$ and $t_c$) across all studied domains (Fig. 5). Here, $t_c$ is the time at which cultural memory overtakes communicate memory, which, according to the model, can be approximated as (see 'Transition time' in the Methods and Supplementary Model):

$$t_c = \frac{1}{p+r-q}\log\left(\frac{(p+r)(p-q)}{rq}\right) \tag{7}$$

Although our results suggest that the functional form of the decay in attention function is universal across multiple cultural domains, its parameters are informative of the domain-specific decay dynamics (Fig. 5). When comparing the obtained parameters, we find that the decay rates of communicative memory are much larger than those of cultural memory ($p \gg q$), as suggested by the literature[2] (Fig. 5a). In addition, we find that communicative memory decays much faster for music and movies than for biographies (Fig. 5c), resulting in critical times that are relatively low for music, movies and papers (5–10 years; Fig. 5d) and much longer for biographies (15–30 years). In other words, for biographies, the era dominated by communicative memory lasts longer than the era dominated by cultural memory.

Together, these results show that the biexponential decay predicted from formalizing the mechanisms suggested by the literature on collective memory provides a universally good approximation for the decay of memory across a wide variety of cultural domains.

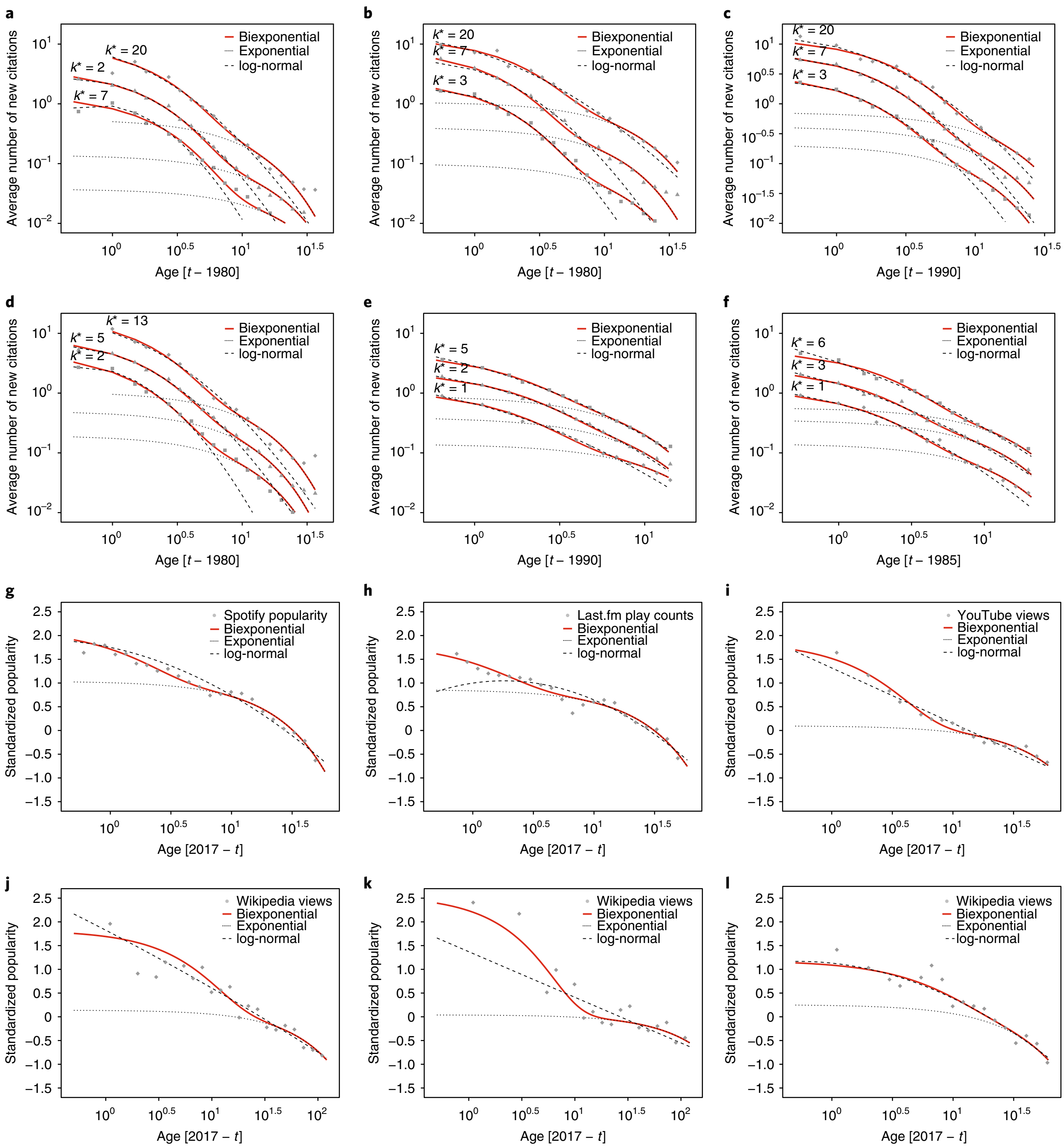


**Fig. 3 | The universal decay of collective memory.** **a**–**f**, Average number of new citations received by: all papers published in *Physical Review B* in 1980 ($n = 1,415$) (**a**), all papers published in *Physical Review D* in 1980 ($n = 803$) (**b**), all papers published in *Physical Review Letters* in 1990 ($n = 1,904$) (**c**), all papers published in *Physical Review Letters* in 1980 ($n = 1,202$) (**d**), all mechanical patents granted in 1990 ($n = 20,296$) (**e**) and all chemical patents granted in 1985 ($n = 14,749$) (**f**). **g**–**l**, For cultural products, we use the standardized levels of online attention for: songs ($n = 18,320$) based on Spotify's popularity index (*y* axis) as a function of the date the song first appeared in the Billboard ranking (*x* axis) (**g**), songs ($n = 15,275$) based on Last.fm's play counts (*y* axis) as a function of the date the song first appeared in the Billboard ranking (*x* axis) (**h**), movies ($n = 14,633$) based on YouTube's view counts (*y* axis) as a function of the date the movie was released (*x* axis) (**i**), tennis players ($n = 624$) based on Wikipedia's page views (*y* axis) as a function of the date that the tennis player was included in the Top 600 International males singles tennis player (*x* axis) (**j**), Olympic medalists ($n = 526$) based on Wikipedia's page views (*y* axis) as a function of the date of the middle of the career of the Olympic medalist (**k**), and basketball players ($n = 592$) based on Wikipedia's page views (*y* axis) as a function of the date that the basketball player starts his career (*x* axis) (**l**). The dashed and dotted lines show the log-normal decay used by Wang et al.[12] and the exponential decay used by Higham et al.[13], respectively.

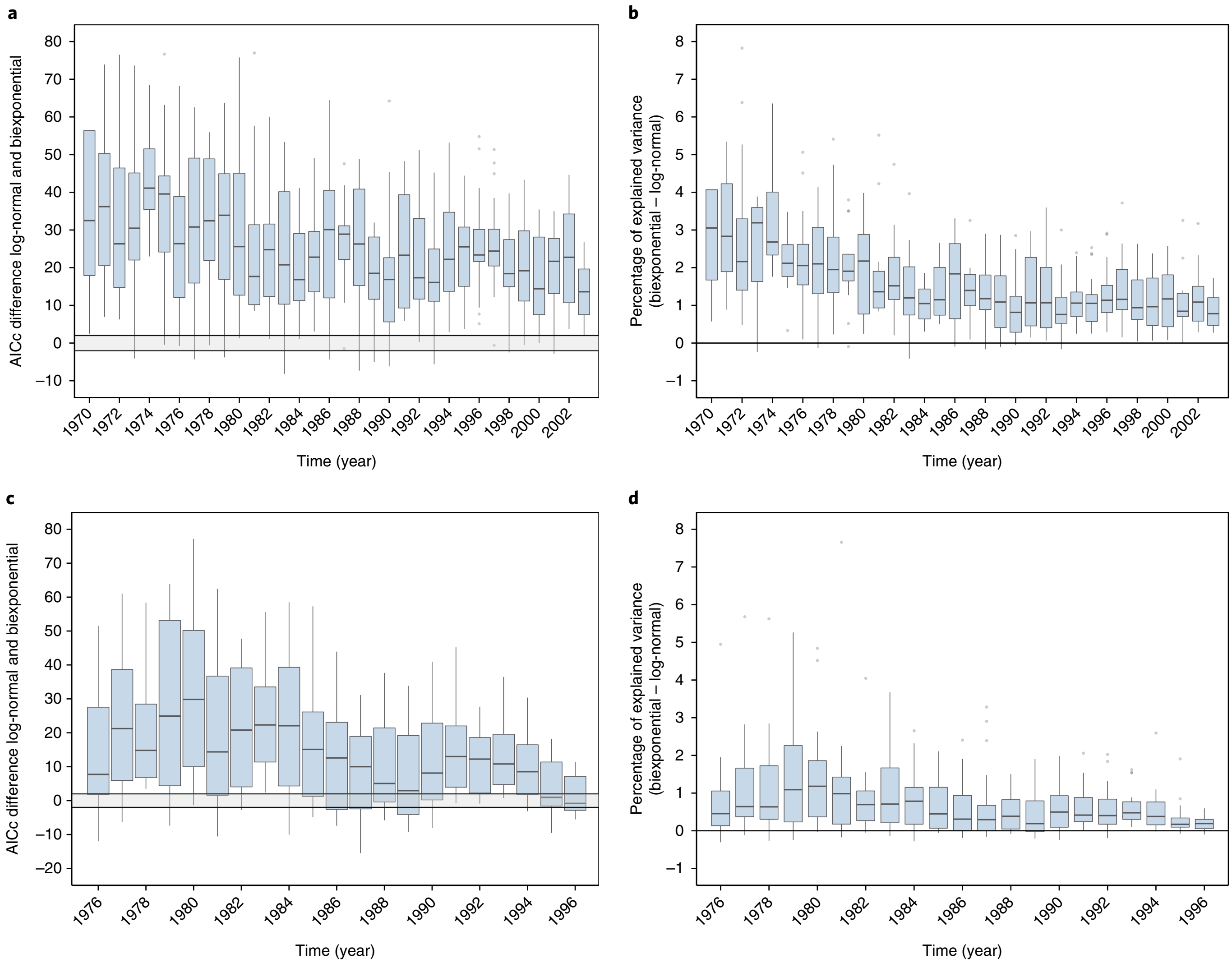


**Fig. 4 | Goodness of fit for all cohorts of APS papers ($n = 485{,}105$) and USPTO patents ($n = 1{,}681{,}690$).** **a**, Difference of the AICc for the log-normal and biexponential decay functions for APS papers. **b**, Difference of the $R^2$ for the biexponential and log-normal decay functions for APS papers. **c**, Difference of the AICc for the log-normal and biexponential decay functions for USPTO patents. **d**, Difference of the $R^2$ for the biexponential and log-normal decay functions for USPTO patents. The grey zones represent the non-significant difference between two models (**a** and **c**). The black lines represent equal goodness of fit (**b** and **d**). Boxplot elements represent individual curves. The lower and upper hinges correspond to the 25th and 75th percentiles respectively. The upper whisker extends from the hinge to the largest value no further than 1.5 times the interquartile distance, between the first and third quartiles. The lower whisker extends from the hinge to the smallest value at most 1.5 times the interquartile distance of the hinge. Data beyond the end of the whiskers, that is, outliers, are plotted individually. We note that the biexponential model outperforms the log-normal model in terms of variance explanation, especially in the long-term description. All of the $R^2$ in **b** and **d** have a $P < 0.001$.

## Discussion

Inspired by Neruda's observation, that love was short and intense, whereas forgetting lingered, we build on the ideas of communicative and cultural memory to show that the decay of the attention received by cultural products and biographies follows a universal decay function that is characterized by two phases: a short-lived and fast-decaying phase connected to communicative memory, and a longer-lived and slower-decaying phase connected to cultural memory. We find that the shape of this function is universal across multiple cultural domains and that its parameters are informative of the attention dynamics that characterize each domain. These findings provide quantitative evidence to validate the concepts of communicative and cultural memory and allow us to better understand how societies forget.

For decades, scholars have been using paper and patent citations to study the spread and adoption of ideas and cultural content[12–14,44,45,50–57]. Indeed, the literature states that the number of citations, A(t), is separable[12,13,15–18] into two mechanisms: (1) the temporal decay, S(t), which captures the time obsolescense and (2) the cumulative citations, c(k), which captures preferential attachment (see 'Decomposition of citing curve' in the Methods and Supplementary Note 4). Yet, although there is consensus on the fact that preferential attachment processes contribute to the spread of cultural products with high levels of attention, there is no consensus on the nature of the functional form capturing the decay of attention. The data show an initially fast decay followed by a milder decline. What gives rise to this unorthodox decay function?

Our results indicate that the fast decay followed by a mild decline observed in these decay functions is a universal biexponential curve that can be derived from a model that builds on two fundamental concepts from the literature on collective memory: communicative memory and cultural memory[1–10]. The agreement between this

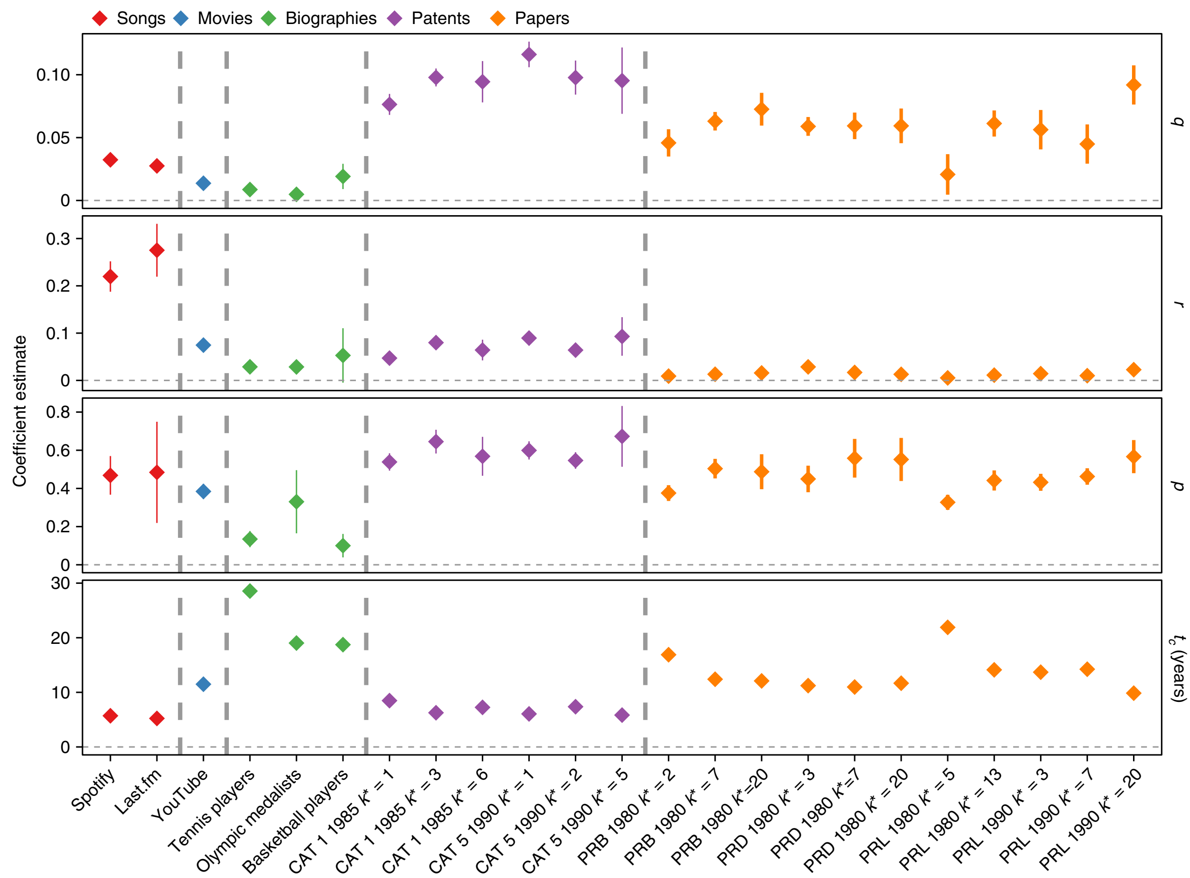


**Fig. 5 | Model parameters described by equation (6) and for the same data deployed in Fig. 2.** Each box corresponds to a model's parameter and the colours represent the type of cultural product. The *y* axis for parameters *q*, *r* and *p* represents the change rate, measured in the number of citations over time (years). The *y* axis for $t_c$ represents the critical time and it is calculated by equation (7) and measured in years. The *x* axis represents the cultural domain analysed. Bars represent the standard deviation of the coefficient estimation (see Supplementary Tables 1–3). PRB, *Physical Review B*; PRD, *Physical Review D*; PRL, *Physical Review Letters*.

model and the empirical data validates these theoretical mechanisms and offers a means to quantify them.

Although the shape of the decay function is universal, its parameters are informative of the decay dynamics of specific systems. For instance, athlete biographies have relatively large critical times compared to songs, movies, papers and patents, meaning that athletes are remembered mainly through oral culture for as long as a couple of decades after their main accomplishment. Conversely, songs have a relatively high rate of transfer ($r$) from communicative to cultural memory and are short lived in communicative memory. In fact, both songs and movies are short lived in communicative memory, but movies live in communicative memory a bit longer than songs, probably because of the high output of the music industry.

But why would communicative memory feed cultural memory? The rationale behind communicative memory feeding cultural memory is that, after each communicative act, the probability that a record is created increases. The parameter $r$ intends to capture, on average, how communicative memory translates into cultural memory. We acknowledge that there are more complex mechanisms associated with this process, for instance, cultural memory should also feed communicative memory. Yet, despite these simplifying assumptions, the model employed here still explains much of the variation observed in the data.

According to our model, in the beginning, most of the attention comes from acts of communication, but this changes over time. Indeed, after a critical time ($t_c$), these cultural products receive more attention from records than from acts of communication. For instance, soon after their release, scientific papers are discussed at conferences, in media, magazines and the news. This generates an excess of attention for newer cultural products and the creation of new records referring to that product. Yet, once the conversation is over, the attention coming from the consultation of these records becomes dominant.

Nevertheless, it is interesting to think about the mechanisms that could contribute to the reduction of communicative memory or the flattening of the biexponential function. For example, the level of coordinated consumption of cultural goods (for example, how much people like to go to the movies together) could modulate how much that cultural good is discussed, and hence, the size of the communicative bump. In addition, exogenous effects, such as the cancellation of a conference owing to the weather could reduce the communicative memory effects for the papers discussed in those conferences.

Our results support the hypotheses that the decay of human collective memory involves the combined decay of communicative and cultural memory and that the decay function is universal across multiple cultural domains. These findings allow us to explain the dynamics of the attention received by scientific papers, patents, songs, movies and biographies during its lifetime, and suggest that the dynamics of human collective memory follow a universal decay function.

## Methods

**Data.** We use two types of data sources: time-series data for scientific papers and patents, and cross-section data for songs, movies and biographies (summarized in Table 1). The American Physical Society (APS) corpus collects data about the attention pattern of physics articles from 12 different journals, between 1896 and 2016. For our analysis, we use a prospective approach (See Supplementary Methods) for all papers published between 1970 and 2003 in *Physical Review Letters* and in *Physical Review A* to *E*[14,53,58] ($n=485{,}105$). The US Patent and Trademark Office (USPTO)[59,60] contains information about patents granted between 1976 and 2005. We use all patents granted between 1976 and 1995 in all categories ($n=1{,}681{,}690$): chemical (CAT 1), computers and computation (CAT 2), drugs and medical (CAT 3), electrical and electronic (CAT 4), mechanical (CAT 5) and others (CAT 6). For both patents and papers, we construct two time series, one for the number of citations obtained in each time window, and another for the accumulated citations obtained up to a given time. Because we are interested in characterizing the dynamics of relative attention, we adjust the time series by normalizing it by the number of papers published in a journal each year[13,14] (see Supplementary Methods).

For songs, movies and online biographies, we use cross-section data, that is, data collected by observing songs, movies and biographies at the same point in time. We use different inclusion criteria—what cultural products are included in our sample—for each type of cultural content. For songs, we use weekly ranking data from the 'Hot-100 Billboard's ranking'[61] between October 1958 and July 2017. To measure online attention, we use Spotify's popularity index[62] (a direct function of play counts) taken on October 2016 and July 2017 ($n=18{,}320$), and last.fm's ($n=15{,}275$) play counts[63] for the last week of July 2017 (see Supplementary Methods). We also collect data on 14,633 movies released between 1937 and 2017 that have obtained more than 1,000 votes in IMDB (https://www.imdb.com) as of July 2017. To measure the current online attention of movies, we use the play counts for the trailer of each movie taken from YouTube (https://www.youtube.com) ($n=14{,}633$). For online biographies, we focus on basketball, tennis and Olympic medal winners. For basketball players, we consider the 'Slam 500 Greatest NBA Players of All Times' ($n=592$), for tennis players, we consider the 'Top 600 International males singles tennis player' ($n=624$) and for Olympic medal winners, we consider athletes who have won more than three gold medals ($n=526$). Current online attention was measured using the number of page views received by the Wikipedia biography (https://en.wikipedia.org) of each athlete between July 2016 and June 2017 (for more information, see Supplementary Methods).

**Decomposition of citing curve.** Mathematically, in our approach, the temporal decay curves describing the number of citations or attention $A(t)$ received by a paper, patent or piece of cultural content (Fig. 1a) can be expressed as a function of two parameters: (1) its age $t$, and (2) the cumulative citations received by that paper, patent or cultural piece $k$. Formally, it has been shown that $A(t)$ is separable[12,13,15–18], as $A(t)=c(k)\times S(t)$, where $c(k)$ captures the effects of preferential attachment (Fig. 1b) and $S(t)$ captures the temporal decay (Fig. 1c).

The solid line (Fig. 1a) shows the average number of citations received by papers published in *Physical Review B* in 1990 ($A(t)$) as a function of their age. $A(t)$ describes the traditional increase and decline known to characterize knowledge diffusion or cultural product adoption curves[13,15–18].

Figure 1b shows the preferential attachment component, by presenting the number of new citations ($\Delta c$) received by a paper as a function of its cumulative citations ($c(k)$)[19,20]. Figure 1c shows the temporal decay component ($S(t)$), representing the number of new citations received by papers with the same number of cumulative citations $k=k^{*}$ as a function of their age; that is, the dashed lines show papers for which the effect of preferential attachment is kept constant: $A(t)|_{k=k^{*}}=c(k^{*})\times S(t)$. Here, we observe the initially fast decay followed by a milder decline.

**Model.** Here, we formalize this intuition by proposing a model for the decay of the attention received by a cultural piece. We took inspiration from collective memory studies and nuclear decay. We solve the model analytically as follow:

$$\frac{du}{dt}=-pu-ru=-(p+r)u \tag{8}$$

$$\frac{dv}{dt}=ru-qv=ru-qv \tag{9}$$

We can write this using matrix representation:

$$\begin{pmatrix}\frac{du}{dt}\\ \frac{dv}{dt}\end{pmatrix}=\begin{pmatrix}-(p+r) & 0\\ r & -q\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}$$

Where the initial conditions are:

$$\begin{pmatrix}u(0)\\ v(0)\end{pmatrix}=\begin{pmatrix}N\\ 0\end{pmatrix}$$

Then, to solve the equation system, we first have to find the eigen values of the $2\times 2$ matrix, by calculatig the matrix determinant (det), this is:

$$\det(A-\lambda I)=0 \tag{10}$$

Solving for $A$, we find $\lambda_1=-(p+r)$ and $\lambda_2=-q$. Now, we have to find the eigen vectors, this is:

$$\begin{pmatrix}-(p+r)-\lambda & 0\\ r & -q-\lambda\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}$$

Using both $\lambda_1=-(p+r)$ and $\lambda_2=-q$, we find that the eigen vectors are:

$$\boldsymbol{\eta}_{\lambda_1}=\left(1,\frac{r}{q-p-r}\right) \tag{11}$$

$$\boldsymbol{\eta}_{\lambda_2}=(0,1) \tag{12}$$

Now, we have that, the general solution is:

$$\mathbf{x}(t)=C_1e^{\lambda_1 t}\boldsymbol{\eta}_{\lambda_1}+C_2e^{\lambda_2 t}\boldsymbol{\eta}_{\lambda_2} \tag{13}$$

Using initial conditions, we find that $C_1=N$ and $C_2=\frac{Nr}{p+r-q}$. Thus:

$$u(t)=Ne^{-(p+r)t} \tag{14}$$

$$v(t)=\frac{Nr}{p+r-q}(e^{-qt}-e^{-(p+r)t}) \tag{15}$$

Finally, the biexponential model is:

$$S(t)=N\left[e^{-(p+r)t}+\frac{r}{p+r-q}(e^{-qt}-e^{-(p+r)t})\right] \tag{16}$$

**Transition time.** An interesting parameter here is the critical time, which is the time when the temporal scale occurs. We calculate the critical time as:

$$\left.\frac{d\log S}{dt}\right|_{t=t_c}=-(1+\delta)q \tag{17}$$

where $\delta\sim 1$.

$$\frac{d\log S}{dt}=\frac{1}{S}\left[-(p+r)e^{-(p+r)t_c}+\frac{r}{p+r-q}(-qe^{-qt_c}+(p+r)e^{-(p+r)t_c})\right] \tag{18}$$

$$\frac{d\log S}{dt}=\frac{1}{S}\left[(p+r)e^{-(p+r)t_c}\left(\frac{r}{p+r-q}-1\right)+\frac{rq}{p+r-q}e^{-qt_c}\right] \tag{19}$$

By definition, $S(t_c)\approx\frac{r}{p+r-q}e^{-qt_c}$. Thus,

$$\left.\frac{d\log S}{dt}\right|_{t=t_c}\approx\frac{(p+r)(p+r-q)}{r}e^{-(p+r-q)}\left(\frac{-p+q}{p+r-q}\right)+q=-(1+\delta)q \tag{20}$$

$$\Rightarrow -(1+\delta)q=\frac{(p+r)}{r}e^{-(p+r-q)}(-p+q)+q \tag{21}$$

$$\Rightarrow q\delta=\frac{(p+r)}{r}e^{-(p+r-q)}(p-q) \tag{22}$$

$$\Rightarrow t_c(\delta)=\frac{1}{p+r-q}\left[\log\left(\frac{(p+r)(p-q)}{rq}\right)-\log\delta\right] \tag{23}$$

In the main text, we have used $\delta=1$, meaning that we have defined the critical time $t_c$ as the time when the decay rate of $S$ is equal to $2q$.

**Model fitting.** We fit our model to paper, patent, song, movie and biography data. In particular, and motivated for accuracy, we fit the logarithm of equation (6), which means that we fit the follow equation:

$$\log(\overline{S(t)})=\log\left[\frac{N}{p+r-q}[(p-q)e^{-(p+r)t}+re^{-qt}]\right] \tag{24}$$

where $\overline{S(t)}$ corresponds to the average of new citations for papers and patents, and $\underline{S}(t) = (S(t) - \overline{Pop}) / \sigma_{Pop}$ corresponds to the standardized current popularity for songs, movies and biographies. $\overline{Pop}$ is the average popularity and $\sigma_{Pop}$ is the standard deviation of current popularity of the decay curves. Those results are shown in the main text for songs (Fig. 3g), movies (Fig. 3i), tennis players (Fig. 3j), Olympic medalists (Fig. 3k) and basketball players (Fig. 3l). Please see the Supplementary Software section for an example anonymized data set and the code used to produce the results.

**Goodness of fit.** We analysed three levels of accomplishment, ($k^*$), for each cohort of APS papers (507 groups of papers) and USPTO patents (480 groups of patents). We compute the AICc (Fig. 4a) to compare the biexponential and log-normal models, corrected by the size of the sample as follow:

$$\mathrm{AICc} = 2k - 2\ \ln(\hat{L}) + \frac{2k^2 + 2k}{n-k-1} \tag{25}$$

where $\hat{L}$ is the maximum value of the likelihood function for the model. In addition, we calculate the $R^2$ as the square of the correlation between the observed value and the predicted value. In Fig. 4a, we observe that the difference for AICc in both papers and patents is significantly bigger than two. It means that, after accounting by the size of the sample and the number of parameters of the model, the biexponential decay presents substantial evidence to be better describing the whole decay (we note that a lower AICc means less information lost in the fit, which is the reason why the difference is positive in the figure). The grey stripe represents the zone where both log-normal and biexponential are equally good at describing the behaviour. We observe that, even after correcting by the size of the sample and by penalizing the number of parameters, the biexponential model offers a more accurate description of the decay function. We also calculate the difference of the adjusted pseudo-$R^2$ (Fig. 4b) between biexponential and log-normal decay. We observe that in both papers and patents the $R^2$ is bigger for biexponential decay than for log-normal decay. We observe that the biexponential model is always better than the log-normal model, especially when it comes to the long-term behaviour of the decay. All models presented in Fig. 5 are summarized in Supplementary Tables 1–3.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Code availability.** The entire analysis, data processing and fitting were done using the standard R libraries (https://www.r-project.org/). You can find the anonymized data on paper citations and the code used to produce the results in this article in zip format in the Supplementary Software section (requires R).

## Data availability

The data sets from the APS, analysed during the current study, are available in the APS Data Sets for Research repository, under request: https://journals.aps.org/datasets. The data sets of the USPTO, analysed during the current study, are available in the NBER repository: http://www.nber.org/patents/. The data sets for songs, movies and biographies generated and analysed during the current study are available from the corresponding authors upon reasonable request.

Received: 18 April 2018; Accepted: 18 October 2018;
Published online: 10 December 2018

## References


1. Halbwachs, M. *La Mémoire Collective* (Albin Michel, Paris, 1997).
2. Assmann, J. Collective memory and cultural identity. *New Ger. Crit.* **65**, 125–133 (1995).
3. Assmann, J. *Das kulturelle Gedächtnis. Schrift, Erinnerung und politische Identität in frühen Hochkulturen* (Verlag C. H. Beck, Munich, 2007).
4. Assmann, J. in *Cultural Memories* (eds Meusburger, P. et al.) 15–27 (Springer, Heidelberg, 2011).
5. Wertsch, J. V. *Voices of Collective Remembering* (Cambridge Univ. Press, Cambridge, 2002).
6. Goldhammer, A., Nora, P. & Kritzman, L. D. *Realms of Memory: The Construction of the French Past* (Columbia Univ. Press, New York, 1998).
7. Roediger, H. & DeSoto, K. Forgetting the presidents. *Science* **346**, 1106–1109 (2014).
8. Roediger, H. L. III, Zaromb, F. M. & Butler, A. C. in *Memory in Mind and Culture* (eds Boyer, P. & Wertsch, J. V.) 138–170 (Cambridge Univ. Press, Cambridge, 2009).
9. Zaromb, F., Butler, A. C., Agarwal, P. K. & Roediger, H. L. III. Collective memories of three wars in United States history in younger and older adults. *Mem. Cognit.* **42**, 383–399 (2014).
10. Rubin, D. C. How quickly we forget. *Science* **346**, 1058–1059 (2014).
11. García-Gavilanes, R., MollgaardA., Tsvetkova, M. & Yasseri, T. The memory remains: understanding collective memory in the digital age. *Sci. Adv.* **3**, e1602368 (2017).
12. Wang, D., Song, C. & Barabási, A.-L. Quantifying long-term scientific impact. *Science* **342**, 127–132 (2013).
13. Higham, K. W., Governale, M., Jaffe, A. B. & Zülicke, U. Fame and obsolescence: disentangling growth and aging dynamics of patent citations. *Phys. Rev. E* **95**, 042309 (2017).
14. Higham, K., Governale, M., Jaffe, A. & Zülicke, U. Unraveling the dynamics of growth, aging and ination for citations to scientific articles from specific research fields. *J. Informetr.* **11**, 1190–1200 (2017).
15. Valverde, S., Solé, R. V., Bedau, M. A. & Packard, N. Topology and evolution of technology innovation networks. *Phys. Rev. E* **76**, 056118 (2007).
16. Csárdi, G., Strandburg, K. J., Zalányi, L., Tobochnik, J. & Érdi, P. Modeling innovation by a kinetic description of the patent citation system. *Physica A* **374**, 783–793 (2007).
17. Golosovsky, M. & Solomon, S. Stochastic dynamical model of a growing citation network based on a self-exciting point process. *Phys. Rev. Lett.* **109**, 098701 (2012).
18. Dorogovtsev, S. N. & Mendes, J. F. F. Evolution of networks with aging of sites. *Phys. Rev. E* **62**, 1842–1845 (2000).
19. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
20. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
21. Price, D. D. S. A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inf. Sci.* **27**, 292–306 (1976).
22. Yule, G. U. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. R. Soc. Lond. B* **213**, 21–87 (1925).
23. Merton, R. K. The Matthew effect in science, II: cumulative advantage and the symbolism of intellectual property. *Isis* **79**, 606–623 (1988).
24. Assmann, J. in *Cultural Memory Studies. An International and Interdisciplinary Handbook* (eds Erll, A. & Nünning, A.) 109–118 (Walter de Gruyter, Berlin, 2008).
25. Hirst, W., Yamashiro, J. K. & Coman, A. Collective memory from a psychological perspective. *Trends. Cogn. Sci.* **22**, 438–451 (2018).
26. Roediger, H. L. & DeSoto, K. A. Recognizing the presidents: was Alexander Hamilton president? *Psychol. Sci.* **27**, 644–650 (2016).
27. Rubin, D. C. *Memory in Oral Traditions: the Cognitive Psychology of Epic, Ballads, and Counting-Out Rhymes* (Oxford Univ. Press, Oxford, 1995).
28. Hammack, P. L. Narrative and the politics of meaning. *Narrat. Inq.* **21**, 311–318 (2011).
29. Sperber, D. & Hirschfeld, L. A. The cognitive foundations of cultural stability and diversity. *Trends Cogn. Sci.* **8**, 40–46 (2004).
30. Buskell, A. What are cultural attractors? *Biol. Philos.* **32**, 377–394 (2017).
31. Richerson, P. J. & Boyd, R. *Not by Genes Alone: How Culture Transformed Human Evolution* (Univ. Chicago Press, Chicago, 2005).
32. Storm, B. C., Bjork, E. L. & Bjork, R. A. On the durability of retrieval-induced forgetting. *J. Cogn. Psychol.* **24**, 617–629 (2012).
33. Garcia-Bajos, E., Migueles, M. & Anderson, M. Script knowledge modulates retrieval-induced forgetting for eyewitness events. *Memory* **17**, 92–103 (2009).
34. Cuc, A., Koppel, J. & Hirst, W. Silence is not golden: a case for socially shared retrieval-induced forgetting. *Psychol. Sci.* **18**, 727–733 (2007).
35. Echterhoff, G., Higgins, E. T. & Levine, J. M. Shared reality: experiencing commonality with others' inner states about the world. *Perspect. Psychol. Sci.* **4**, 496–521 (2009).
36. Coman, A. & Hirst, W. Social identity and socially shared retrieval-induced forgetting: the effects of group membership. *J. Exp. Psychol. Gen.* **144**, 717–722 (2015).
37. Coman, A., Stone, C. B., Castano, E. & Hirst, W. Justifying atrocities: the effect of moral-disengagement strategies on socially shared retrieval-induced forgetting. *Psychol. Sci.* **25**, 1281–1285 (2014).
38. Stone, C. B., Barnier, A. J., Sutton, J. & Hirst, W. Building consensus about the past: schema consistency and convergence in socially shared retrieval-induced forgetting. *Memory* **18**, 170–184 (2010).
39. Kurzban, R., Tooby, J. & Cosmides, L. Can race be erased? Coalitional computation and social categorization. *Proc. Natl Acad. Sci. USA* **98**, 15387–15392 (2001).
40. Yu, A. Z., Ronen, S., Hu, K., Lu, T. & Hidalgo, C. A. Pantheon 1.0, a manually verified dataset of globally famous biographies. *Sci. Data* **3**, 150075 (2016).
41. Jara-Figueroa, C., Yu, A. Z. & Hidalgo, C. A. How the medium shapes the message: printing and the rise of the arts and sciences. *PLoS ONE* (in the press).
42. Skiena, S. & Ward, C. B. *Who's Bigger? Where Historical Figures Really Rank* (Cambridge Univ. Press, Cambridge, 2014).
43. Kanhabua, N., Nguyen, T. N. & Niederée, C. What triggers human remembering of events? A large-scale analysis of catalysts for collective memory in Wikipedia. In *Proc. 14th ACM/IEEE-CS Joint Conference on Digital Libraries* 341–350 (IEEE, 2014).
44. Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. Atypical combinations and scientific impact. *Science* **342**, 468–472 (2013).
45. Mukherjee, S., Romero, D. M., Jones, B. & Uzzi, B. The nearly universal link between the age of past knowledge and tomorrow's breakthroughs in science and technology: the hotspot. *Sci. Adv.* **3**, e1601315 (2017).

46. Ronen, S. et al. Links that speak: the global language network and its association with global fame. *Proc. Natl Acad. Sci. USA* **111**, E5616–E5622 (2014).
47. Ferron, M. & Massa, P. Beyond the encyclopedia: collective memories in Wikipedia. *Mem. Stud.* **7**, 22–45 (2014).
48. Yucesoy, B. & Barabási, A.-L. Untangling performance from success. *EPJ Data Sci.* **5**, 17 (2016).
49. Halbwachs, M. *On Collective Memory* (Univ. Chicago Press, Chicago, 1992).
50. Fortunato, S. et al. Science of science. *Science* **359**, eaao0185 (2018).
51. Radicchi, F., Fortunato, S. & Castellano, C. Universality of citation distributions: toward an objective measure of scientific impact. *Proc. Natl Acad. Sci. USA* **105**, 17268–17272 (2008).
52. Kuhn, T., Perc, M. & Helbing, D. Inheritance patterns in citation networks reveal scientific memes. *Phys. Rev. X* **4**, 041036 (2014).
53. Sinatra, R., Wang, D., Deville, P., Song, C. & Barabasi, A.-L. Quantifying the evolution of individual scientific impact. *Science* **354**, aaf5239 (2016).
54. King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J. & West, J. D. Men set their own cites high: gender and self-citation across fields and over time. *Socius* **3**, 1–22 (2017).
55. Chu, J. S. G. & Evans, J. A. Too many papers? Slowed canonical progress in large fields of science. Preprint at SocArXiv https://doi.org/10.31235/osf.io/jk63c (2018).
56. Yook, S.-H., Radicchi, F. & Meyer-Ortmanns, H. Self-similar scale-free networks and disassortativity. *Phys. Rev. E* **72**, 045105 (2005).
57. Mukherjee, S., Uzzi, B., Jones, B. F. & Stringer, M. in *Knowledge and Networks* (eds Glückler, J. et al.) 243–267 (Springer, Cham, 2017).
58. Shen, H.-W. & Barabasi, A.-L. Collective credit allocation in science. *Proc. Natl Acad. Sci. USA* **111**, 12325–12330 (2014).
59. Hall, B. H., Jaffe, A. B. & Trajtenberg, M. *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools* (National Bureau of Economic Research, 2001); www.nber.org/papers/w8498
60. Jaffe, A. B., Trajtenberg, M. & Henderson, R. Geographic localization of knowledge spillovers as evidenced by patent citations. *Q. J. Econ.* **108**, 577–598 (1993).
61. The Hot 100 Ranking *Billboard* https://www.billboard.com/charts/hot-100 (2016).
62. Spotify web API *Spotify* https://developer.spotify.com/documentation/web-api/ (2017).
63. Last.fm web services *last.fm* https://www.last.fm/api (2017).



## Acknowledgements

C.C. and C.R.-S. acknowledge financial support from Centro de Investigación en Complejidad Social and Universidad del Desarrollo. C.J.-F. and C.A.H. acknowledge support from the MIT Media Lab Consortia. The authors thank F. Pinheiro, T. Roukny, G. Castro-Dominguez, the Centro de Investigación en Complejidad Social, the Collective Learning Group at the MIT Media Lab and the Center for Complex Network Research at Northeastern University for the helpful insights and discussions. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.


## Author contributions

C.C., C.A.H. and A.-L.B. contributed to the study conception and design, interpretation of data and drafting of the manuscript. C.C. and C.J.-F. contributed to the acquisition of data, data analysis, modelling and drafting of the manuscript. C.R.-S. contributed to study conception and design, and interpretation of data.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-018-0474-5.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to C.C. or C.A.H.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

naturerESEARCH

Corresponding author(s): Cesar A. Hidalgo

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size ($n$) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's $d$, Pearson's $r$), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | For data on Papers we asked to American Physical Society for the complete corpus in September 2017. For data on patents, we used the NBER data set publicly available in http://www.nber.org/patents/. For data on Songs, Movies, and biographies we use python 2.7 and we connect to Spotify and LastFm API for songs and to Wikipedia API (https://wikimedia.org/api) for Movies and Biographies. Also, we got all historic the data from Hot-100 billboard ranking from https://www.billboard.com/charts/hot-100 . For movies, we also downloaded data from IMDB from https://www.imdb.com/interfaces/ . |
| Data analysis | We use Python 2.7 for preprocessing data. Then, we use R 3.1.1 for data processing and statistical analysis. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

April 2018

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets from the American Physical Society, analyzed during the current study is available in the APS Data Sets for Research repository, under request https://journals.aps.org/datasets.

The datasets of United State Patents analyzed during the current study is available in the NBER repository, http://www.nber.org/patents/.

The datasets for Songs, Movies, and Biographies generated during and analyzed during the current study are available from the corresponding authors on reasonable request.

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

**Study description**

This is a study of collective behavior. We use time series data and also cross-section data.

**Research sample**

We use all temporal citation data available from UPSTO in NBER repository. We also use all temporal citation data from APS for 6 of its journals (PRA, PRB, PRC, PRD, PRE, and PRL). For songs, movies and biographies, we use cross-section data. Here we use a sample of songs movies and biographies (people) that have a similar accomplishment (reach the billboard ranking at least once, to have more than 1,000 votes in IMDB, and be part of an international performance ranking respectively), in order to be compared accounting by preferential attachment (popularity) effects.

**Sampling strategy**

For patents and papers we don't do samples. For music, movies, and biographies we select all the cultural pieces that have similar accomplishments. For basketball players, we consider the “Slam 500 Greatest NBA Players of All Times,” for tennis players we consider the “Top 600 International males singles tennis player,” and for Olympic medal winners we consider athletes who have won more than three gold medals. We didn't do sample calculation because we prove with time series data for all data of patents and papers that cultural pieces with similar accomplishment (number of cumulated citations) show the same behavior, and because we select all the data for every kind of accomplishment.

**Data collection**

For data on Papers we asked to American Physical Society for the complete corpus in September 2017. For data on patents, we used the NBER data set publicly available in http://www.nber.org/patents/. For data on Songs, Movies, and biographies we use python 2.7 and we connect to Spotify and LastFm API for songs and to Wikipedia API (https://wikimedia.org/api) for Movies and Biographies. Also, we got all historic the data from Hot-100 billboard ranking from https://www.billboard.com/charts/hot-100 . For movies, we also downloaded data from IMDB from https://www.imdb.com/interfaces/ .

**Timing**

We use data for all papers published between 1970 and 2003 in Physical Review Letters (PRL), and in Physical Review A to E.
For The United States Patent and Trademark Office (USPTO) data, we use all patents granted between 1976 and 1995 in all categories.
For songs, we use weekly ranking data from the "Hot-100 Billboard's ranking" between October 1958 and July 2017. To measure online attention, we use Spotify's popularity index taken on October 2016 and July 2017, and last.fm's play counts for the last week of July 2017.
For movies, we collect data on 14,633 movies released between 1937 and 2017 that have obtained more than 1,000 votes in the Internet Movie Database as of July 2017. To measure the current popularity of movies we use the play counts for the trailer of each movie taken from YouTube.
For online biographies we focus on basketball, tennis, and Olympic medal winners. Current popularity was measured using the number of pageviews received by the Wikipedia biography of each athlete between July 2016 and June 2017

**Data exclusions**

There is no data exclusion in this study.

**Non-participation**

There are no participants in this study, therefore, there is no dropout.

**Randomization**

There is no randomization in this study

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
| --- | --- |
| ☒ | ☐ Unique biological materials |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |

### Methods

| n/a | Involved in the study |
| --- | --- |
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |