The Universal Decay of Collective Memory and Attention

Cristian Candia, C. Jara-Figueroa, Carlos Rodriguez-Sickert, Albert-László Barabási, and César A. Hidalgo

Presentation by Olivia Lundelius and Mara Schwartz

Contents of Talk

The Universal Decay of Collective Memory and Attention

Modeling the Temporal Decay of Human Attention

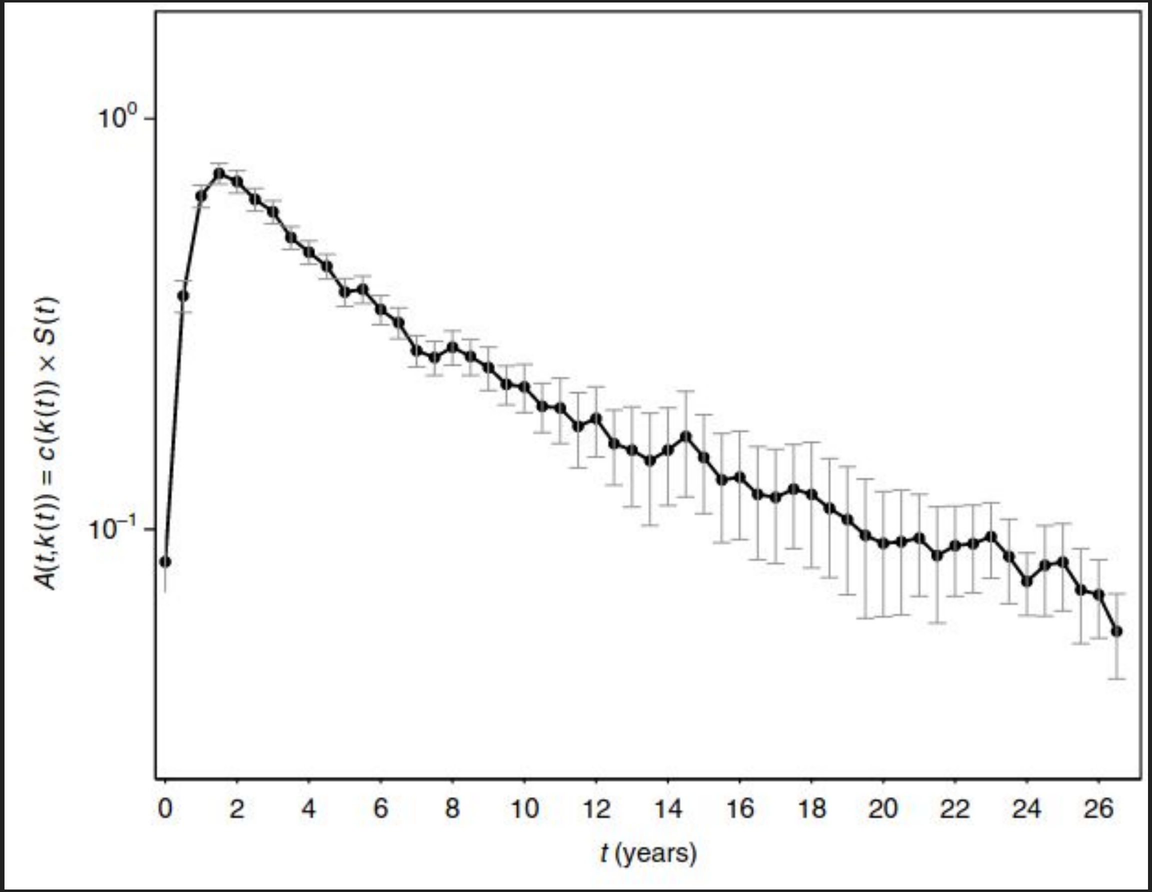
- Does human society experience two phases of memory?
  - Initial phase of high attention
  - Longer and slower phase of forgetting
- Collective memory and attention is comprised of:
  - Communicative Memory
  - Cultural Memory

Knowledge Diffusion Models

- There are few quantitative models on theory of memory and attention
- Focus on two processes:
  - Preferential Attachment (Cumulative Advantage)
    - Attention brings more attention
    - Frequently cited papers are more likely to receive further citations
  - Temporal Decay
    - The decay in attention over time

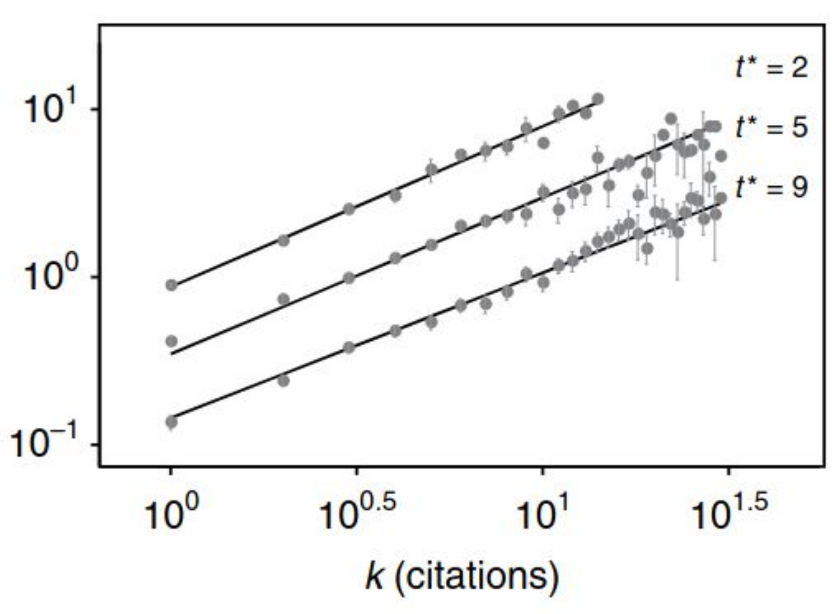
Knowledge Diffusion

Average number of citations received each semester by papers published in *Physical Review B*



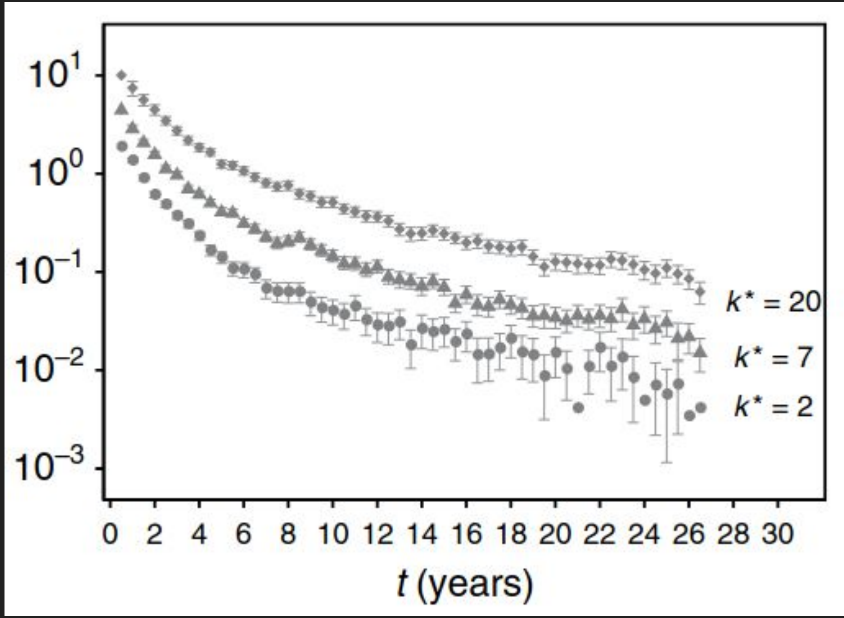
Preferential Attachment

Avg number new citations of papers vs cumulative citations
(grouped by age of paper)



Temporal Decay

Avg number new citations of papers vs age
(grouped by cumulative citations)



Knowledge Diffusion Models (cont.)

- Current describe decay of attention using
  - Exponential functions
  - Log-normal functions
- No consensus on shape of decay function across domains
- The paper proposes biexponential decay function
  - Uses both communicative and cultural memory
  - Statistically better at explaining decay of attention than exponential and log-normal methods
  - Model hopes to capture universality of decay of human collective memory

Hypothesis:

"...The decay of the attention received by cultural products involves the decay of both communicative and cultural memory."

Contents of Talk

The Universal Decay of Collective Memory and Attention

How do *psychologists* approach memory formation and retention?

- Top-Down approaches
  - Familiarity increases memorability of events
  - Narrative templates can shape memories
  - Cultural attractors can increase preservation of memories across generations
- Bottom-Up approaches
  - How micro-level psychological processes can shape social outcomes
  - Forgetting induced through selective retrieval of events
  - People create shared realities with member of their own social group

How do *computational social scientists* approach memory formation and retention?

- Measures byproducts of collective memory and attention rather than direct measures
  - Focus on how collective memory is expressed through content consumption, webpage views, paper and patent citations
  - Topics of increased interest and discussion among people are more likely to generate searches
  - Increases in people’s consultation of related data sources

What are the benefits of the computational social scientist's approach?

- Closer to two-part definition of collective memory
  - Potentiality
    - Existence of record (like a book in a library)
    - Corresponding to cultural memory
  - Actuality
    - Attention received by record (frequency a resource is consulted on a topic)
    - Corresponding to communicative memory

How are we using these concepts in this paper?

- Use of computational social science strand of literature in understanding of memory and attention
    - Due to the computational nature of the experiment
    - Related to psychological method of approaching collective memory and attention
    - Study of selective retrieval
- Approach modeling from a bottom-up perspective
- Can search for a universal average decay function by using data from multiple sources and domains across human culture

Contents of Talk

The Universal Decay of
Collective Memory and Attention

1. Data

- Time-series data
  - Scientific papers
    1970-2003 *Physical Review Letters* and *Physical Review A* to *E*
  - Patents
    1976-1995 USPTO
- Cross-section data
  - Songs
    Oct 1958 - July 2017 weekly Hot-100 Billboard's ranking
  - Movies
    14,633 movies from 1937-2017 with more than 1000 IMDB votes
  - Biographies
    Basketball (Slam 500 Greatest NBA Players of All Times)
    Tennis (Top 600 International males singles tennis player)
    Olympic medal winners (at least 3 gold medals)

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

2. Model

Have cultural and communicative memory coexist, but decay at different rates

Definitions:

$u(t)$: communicative memory at time t

$v(t)$: cultural memory at time t

$S(t)$: current total attention at time t

Assume rate at which communicative memory feeds into cultural memory to be r

Assume communicative memory decay to be $p+r$

Assume cultural memory decay to be q

We can represent this information thus far in the three equations here:

$$S(t) = u(t) + v(t)$$

$$v(t+1) = (1-q)v(t) + ru(t)$$

$$u(t+1) = (1-p)u(t) - ru(t)$$

result is the following system of differential equations:

$$S(t) = u(t) + v(t) \qquad \frac{du}{dt} = -(p+r)u \qquad \frac{dv}{dt} = -qv + ru$$

whose solution is the biexponential function, with $u(0)=N$ and $v(0)=0$:

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

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

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

This is the model that we hope will be an accurate fit for our data

3. Transition Time

t_c: critical time

- The time when cultural products receive more attention from physical records than from acts of communication
- In this paper, defined as the time when the decay rate of S is equal to $2q$

The authors find the critical time through integration to be:

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

4. Model Fitting

The biexponential model was fit to scientific paper, patent, song, movie, and biography data.

For accuracy and display purposes, the authors used the logarithm of the biexponential model to fit the data:

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

$\overline{S(t)}$ being the average of new citations for patents and papers, and the standardized current popularity for movies, songs, and biographies.

Contents of Talk

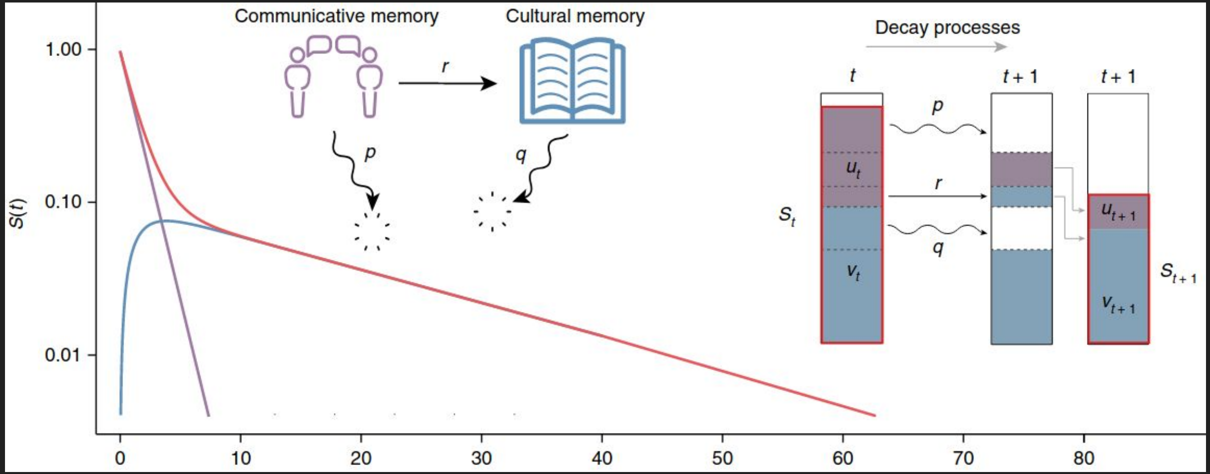
The Universal Decay of Collective Memory and Attention

Scheme of the collective memory model

y-axis is the normalized current level of attention received by a group of comparable cultural pieces
x-axis represents the age of the pieces

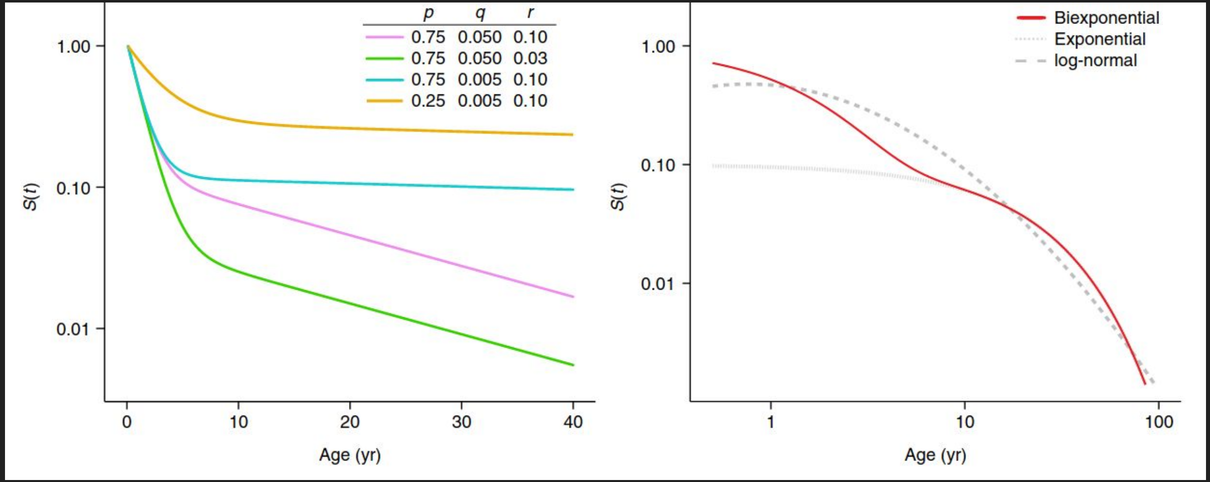
Red curve shows biexponential function predicted by the model in log-lin scale

Blue and purple show 2 exponentials of communicative and cultural memory



Left: the biexponential model for various parameters *p*, *q*, and *r*

Right: Comparison between biexponential model and exponential and log-normal models in log-log scale



Average number of new citations received by:

A all papers in Physical Review B in 1980 (n = 1, 415)

B all papers in Physical Review D in 1980 (n = 803)

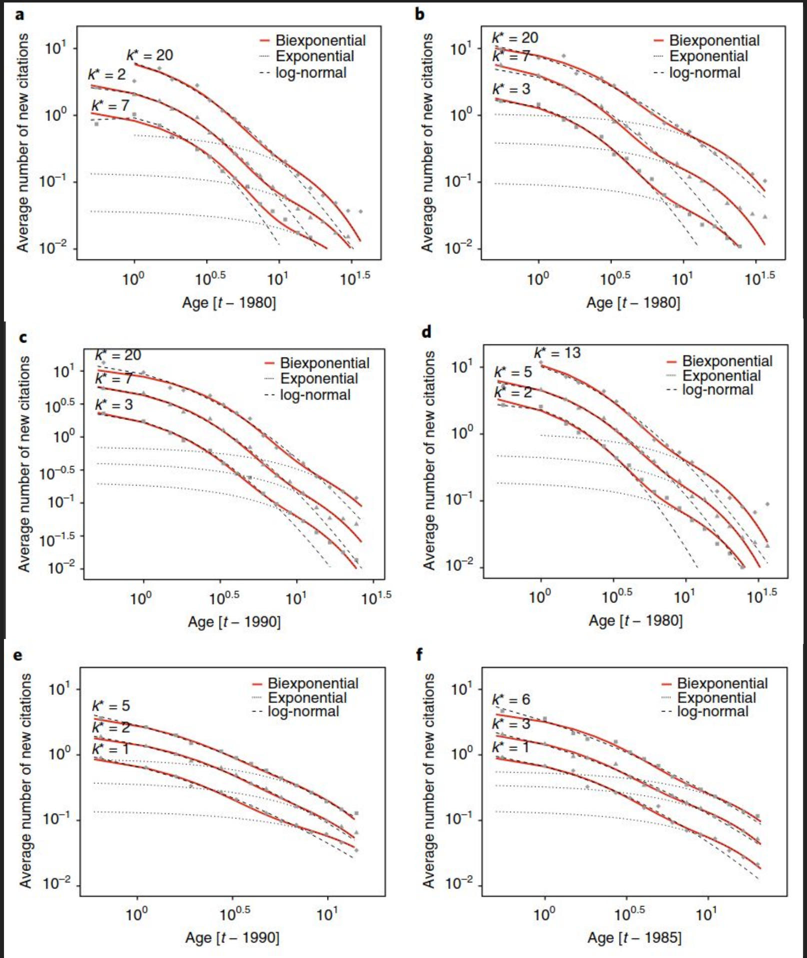
C all papers in Physical Review Letters in 1990 (n = 1, 904)

D all papers in Physical Review L in 1980 (n = 1, 202)

E all Mechanical patents in 1990 (n = 20, 296)

F all Chemical patents in 1985 (n = 14, 749)

Red lines show bi-exponential model fit
Dotted lines show the exponential decay
Dashed lines show the log-normal decay



For cultural products we use the standardized levels of online attention for:

G songs (n = 18, 320) based on Spotify's popularity index versus date first appeared in the Billboard ranking

H songs (n = 15, 275) based on Last.fm's play counts versus date first appeared in the Billboard ranking

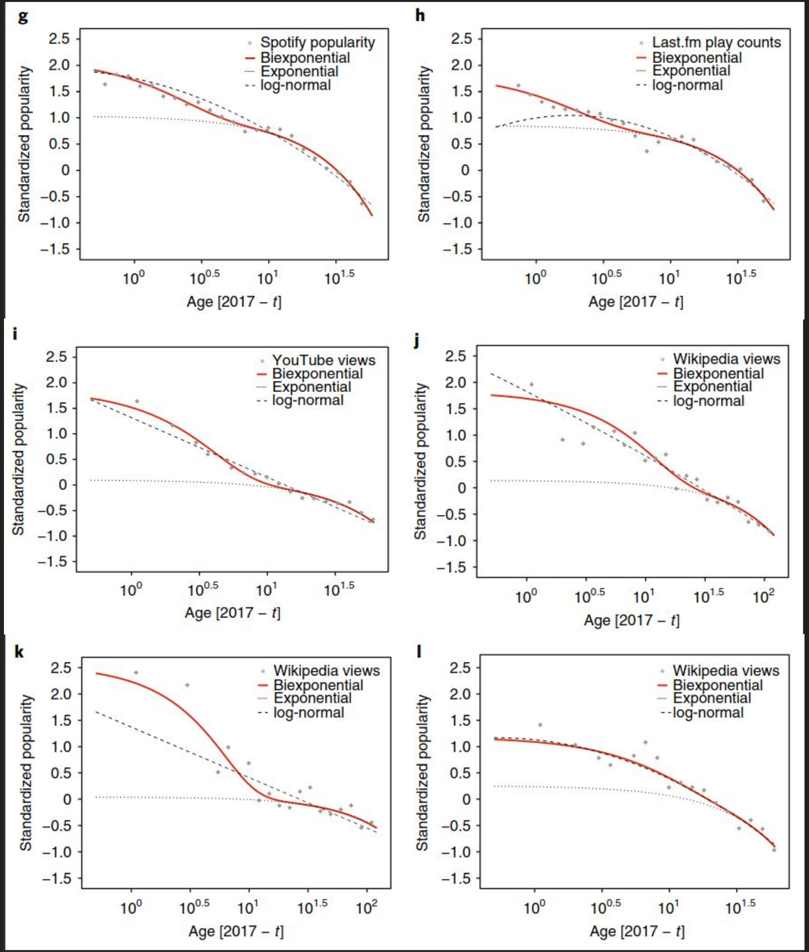
I movies (n = 14, 633) based on YouTube's view counts as a function of the date the movie was released

J tennis players (n = 624) based on Wikipedia's page views versus date that the tennis player was included in the Top 600 International males singles tennis player

K olympic medalist (n = 526) based on Wikipedia's page views versus date of the middle of the medalist's career

L basketball players (n = 592) based on Wikipedia's page views versus date that the Basketball player starts his career

Red lines show bi-exponential model fit
Dotted lines show the exponential decay
Dashed lines show the log-normal decay



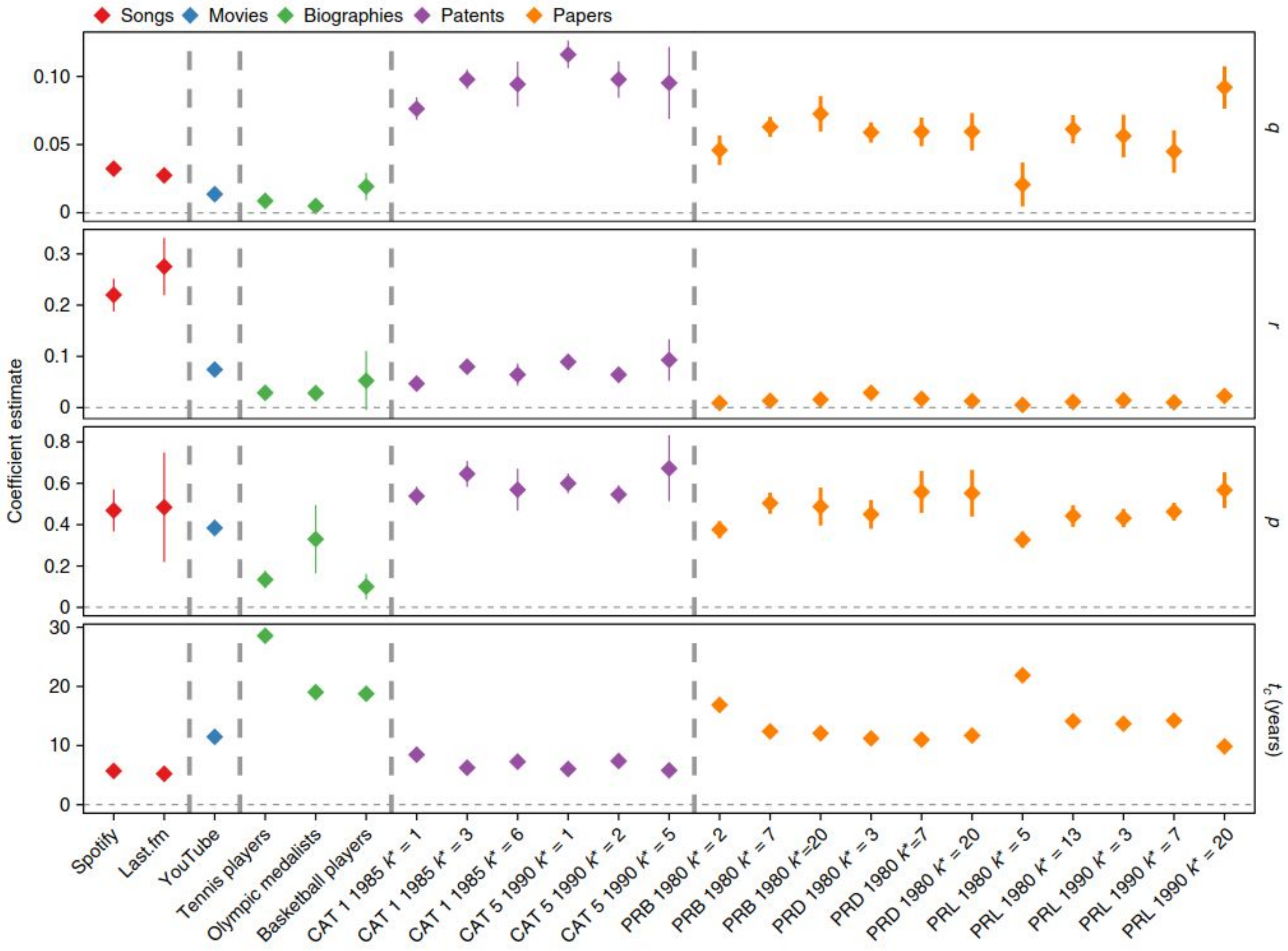
Model Parameters

Each box corresponds to a model's parameter

Colours represent the type of cultural product

y-axis for *q*, *r,* and *p* represent change rate (measured in number of citations over time)

y-axis for t_c represents the critical time in years



Goodness of Fit

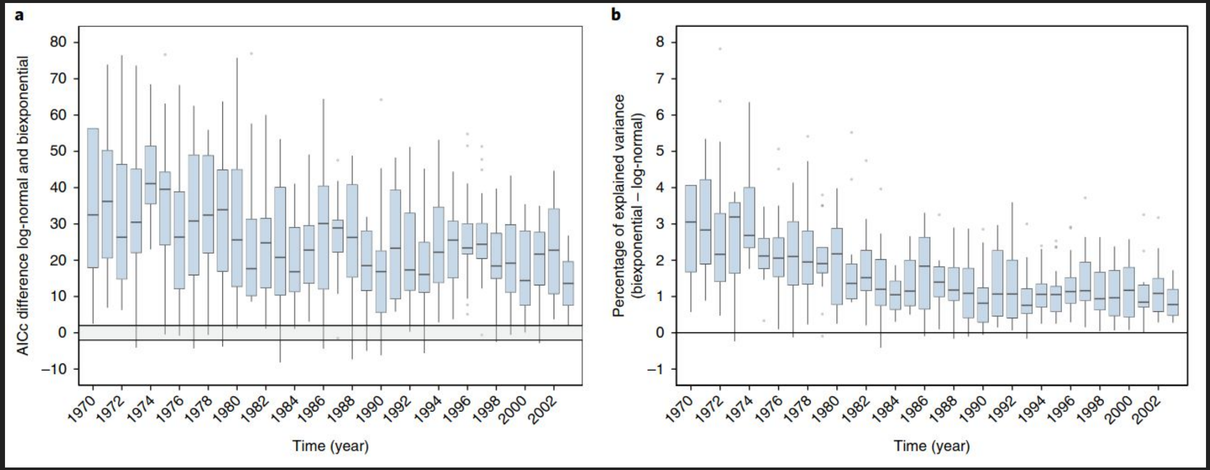
- Analysed AICc to compare the biexponential and log-normal models corrected by the size of the sample
- Calculate R^2 as the square of the correlation between observed and predicted
- Observe that the biexponential decay has substantial evidence to be better at describing the whole decay (lower AICc means less loss)
- Biexponential model consistently better than log-normal model, especially in long-term decay behavior

Difference of the AICc for the log-normal and biexpoential decay functions for APS papers

Difference of the R^2

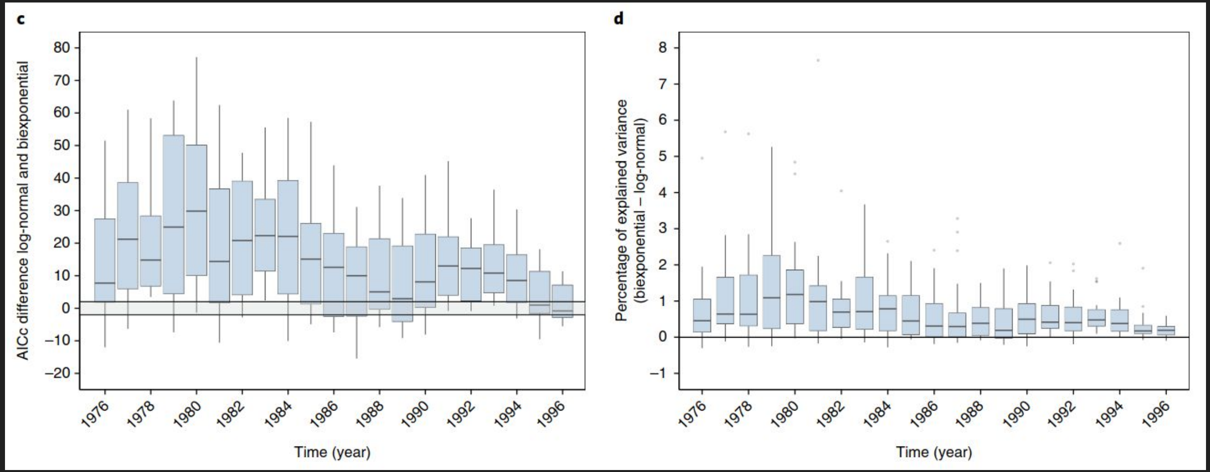
AICc is the information score of the model (the smaller the value, the better the model fit)

R^2 is the goodness of fit metric



Difference of the AICc for the log-normal and biexponential decay functions for USPTO patents

Difference of the R^2



Goodness of fit for APS papers and USPTO patents

What do we take from these graphs?

- Biexponential model provides a more accurate fit to the data than the log-normal and exponential models
- Captures faster decay of communicative memory followed by slower decay of cultural memory
- Reinforces that communicative memory decay is greater than cultural memory decay

“These results show that the biexponential decay... provides a universally good approximation for the decay of memory across a wide variety of cultural domains.”

Contents of Talk

The Universal Decay of Collective Memory and Attention

Discussion

- The decay of attention received by cultural products follows a two-phase universal decay function
    - Fast-decaying phase linked to communicative memory
    - Slower-decaying phase linked to cultural memory
    - The shape is universal across multiple cultural domains
- These findings allow us to better understand how societies forget
    - Provide quantitative evidence to validate the concepts of communicative and cultural memory

Discussion (cont.)

- Shape of biexponential function is universal
- Parameters differ for different domains
- Communicative memory feeds cultural memory
  - Probability a record is created increases after each communication
  - Simplified model using linear r still accurate despite variation
- Cultural products receive more attention from records than communication after critical time
- Results support the hypothesis that decay of human collective memory involves the combined decay of communicative and cultural memory

Questions?

Candia, C., Jara-Figueroa, C., Rodriguez-Sickert, C. et al. The universal decay of collective memory and attention. Nat Hum Behav 3, 82–91 (2019). https://doi.org/10.1038/s41562-018-0474-5