EvoText: A New Text Analysis Tool for the History and Philosophy of Biology

Charles Pence
Louisiana State University
www.charlespence.net

Grant Ramsey
KU Leuven
www.theramseylab.org

1. The origins and goals of evoText
2. The architecture of evoText
3. How to use evoText
4. Preliminary results of evoText

1. The origins and goals of evoText

1. The origins and goals of evoText

What is fitness all about?

www.evoText.org

2. The architecture of evoText

What question do you want to answer?

e.g.: How has the frequency of a term changed over time?

What data do you want to use?

e.g.: Articles in *Nature* from the 1990s

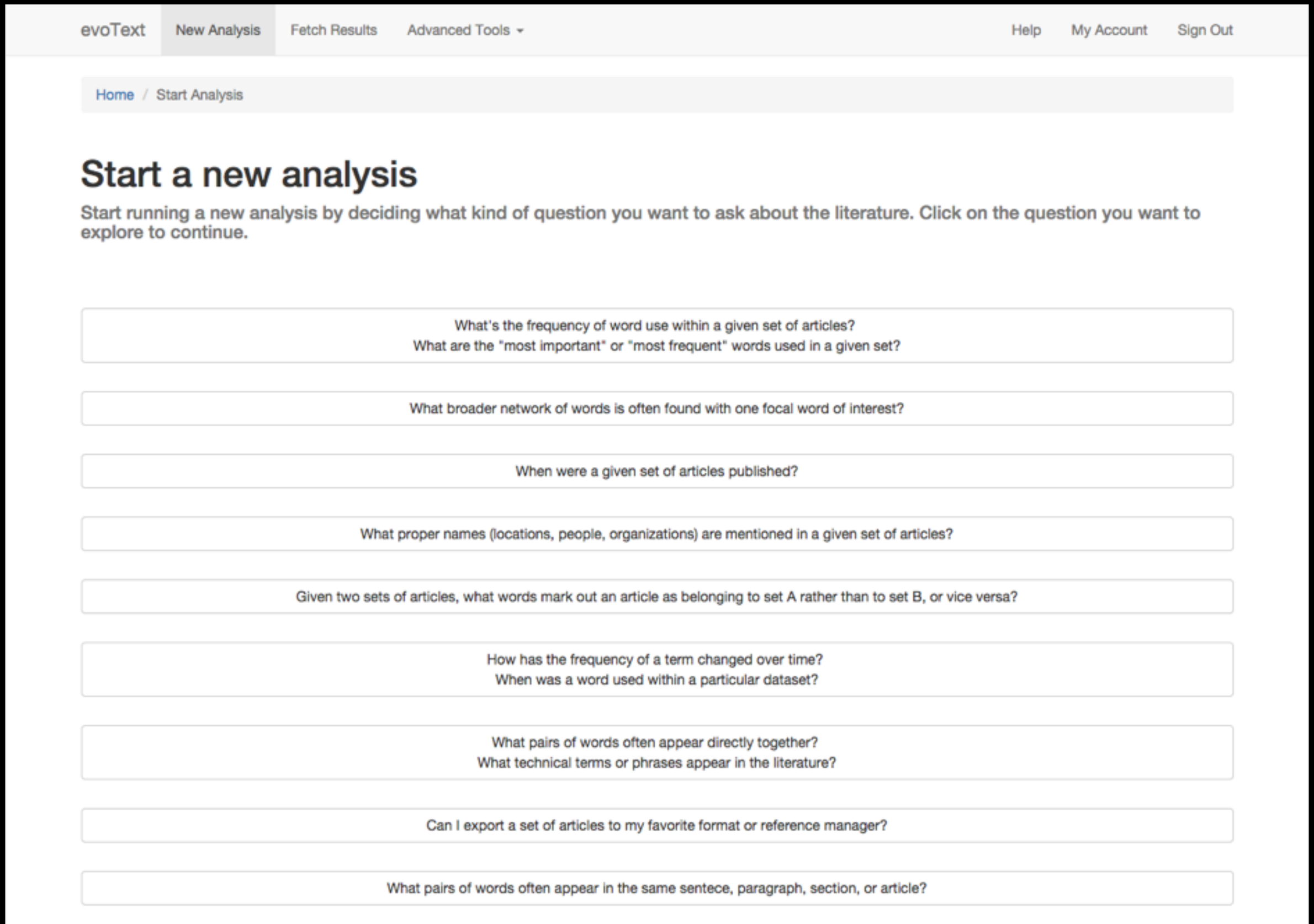
Provide a few more parameters

e.g.: Look for the term “evolution”

(wait for a while)

Get your results!

3. How to use evoText



Plot the use of a term by date

This analysis job takes a single dataset and plots the occurrences of a term within it by date.

This job results in a graph of the occurrences of the given term within your dataset, plotted by date, as well as those occurrences downloadable as a CSV file. This allows us to answer a wide variety of questions:

How has the frequency of use of a term changed over time? *(Input: a dataset of interest, plotting for the use of a given term)*

When was a term first introduced into the literature? *(Input: a dataset of interest, looking for the place when the term is first introduced)*

How has a term moved through the literature? *(Input: comparing these graphs for the same term across different journals and time periods)*

Back Start

Collect data

You now need to determine which datasets this analysis job will run on.

In order to run, this analysis job requires that you provide 1 dataset. Datasets are created by performing searches, then saving a set of search results.

1 dataset still needs to be added

Datasets for this job:

Dataset Name
No datasets specified

Create another dataset | Link an already created dataset

Search for articles... Sort Save Results

449846 articles in database No filters active

« 1 2 3 4 ... 44984 44985 »

The Circadian Neuropeptide PDF Signals Preferentially through a Specific Adenylate Cyclase Isoform AC3 in M Pacemakers of Drosophila More

Laura B. Duvall, Paul H. Taghert | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001337

Brian: The Typographical Error that Brought Early Career Neuroscientists and Artists Together More

Megan J. Dowie, Erin Forsyth, Leah Forsyth | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001340

The Origin of Glucocorticoid Hormone Oscillations More

Jamie J. Walker, Francesca Spiga, Eleanor Waite, Zidong Zhao, Yvonne Kershaw, John R. Terry, Stafford L. Lightman | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001341

The DEAD-Box Protein Dhh1 Promotes Decapping by Slowing Ribosome Movement More

Thomas Sweet, Carrie Kovalak, Jeff Coller | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001342

The Limits to Sustainability Science: Ecological Constraints or Endless Innovation? More

Georgina M. Mace | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001343

The Shifting Boundaries of Sustainability Science: Are We Doomed Yet? More

John H. Matthews, Frederick Boltz | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001344

The Macroecology of Sustainability More

Joseph R. Burger, Craig D. Allen, James H. Brown, William R. Burnside, Ana D. Davidson, Trevor S. Fristoe, Marcus J. Hamilton, Norman Mercado-Silva, Jeffrey C. Nekola, Jordan G. Okie, Wenyun Zuo | *PLoS Biology,* Vol. 10, No. 6 (2012), pp. e1001345

Filter search

Authors

Jr. 1841

R. L. 405

W. F. DENNING 252

G. A. J. C. 177

E. RAY LANKESTER 169

Journal

Nature 383783

Ecology 13754

American Naturalist 11731

Journal of Mammalogy 9222

Evolution 7927

Publication Date

1960–1969 58869

2000–2009 51397

1980–1989 46518

1970–1979 45582

Search for articles... Sort Save Results

32473 articles found Remove All Journal: Nature Year: 1990–1999

« 1 2 3 4 ... 3247 3248 »

You looking at me? More

| *Nature,* Vol. 397, No. 6717 (1999), pp. 313-313

News in Brief More

| *Nature,* Vol. 399, No. 6738 (1999), pp. 724-725

Discovery of the acoustic Faraday effect in superfluid ³He-B More

Y. Lee, T. M. Haard, W. P. Halperin, J. A. Sauls | *Nature,* Vol. 400, No. 6743 (1999), pp. 431-433

Angiogenesis inhibited by drinking tea More

Yihai Cao, Renhai Cao | *Nature,* Vol. 398, No. 6726 (1999), pp. 381-381

Cracking anaerobic bacteria More

| *Nature,* Vol. 401, No. 6750 (1999), pp. 217-218

Measurement of gravitational acceleration by dropping atoms More

Achim Peters, Keng Yeow Chung, Steven Chu | *Nature,* Vol. 400, No. 6747 (1999), pp. 849-852

Science in culture More

| *Nature,* Vol. 401, No. 6755 (1999), pp. 744-744

Arima ascendant More

Filter search

Authors

Alison Abbott 58

David Swinbanks 54

Colin Macilwain 44

Declan Butler 41

David Dickson 35

Advanced search

Search for articles...

32473 articles found Remove All

Create a new dataset

Give these search results a name, and they will be preserved as a new dataset.

Name

Nature '90s

Create Dataset

You looking at me?

| *Nature*, Vol. 397, No. 6717 (1999), pp. 313-313

More

News in Brief

| *Nature*, Vol. 399, No. 6738 (1999), pp. 724-725

More

Discovery of the acoustic Faraday effect in superfluid ³He-B

Y. Lee, T. M. Haard, W. P. Halperin, J. A. Sauls | *Nature*, Vol. 400, No. 6743 (1999), pp. 431-433

More

Angiogenesis inhibited by drinking tea

More

Filter search

Authors

Alison Abbott 88

David Swinbanks 64

Colin Macilwain 44

Declan Butler 41

David Dickson 35

Advanced search

Collect data

All data has now been collected. Click the button to start the analysis!

Set Job Options

Datasets for this job:

Dataset Name	
Nature '90s	Remove

Job options

This job has some more options that you can configure before you get started.

Term frequency date analysis options

Focal word

evolution

Start analysis job Cancel

Fetch analysis results

From here, you can retrieve the data produced by your analyses. Make sure to save it soon, because it is only preserved for 14 days.

Pending analysis tasks

Analysis Task	Dataset	Task Progress
Plot word occurrences by date	Nature '90s	14%: Querying term frequency counts...

Are these tasks taking too long to finish? We can try to terminate all pending jobs, though that might not work. If it fails, e-mail the site administrators.

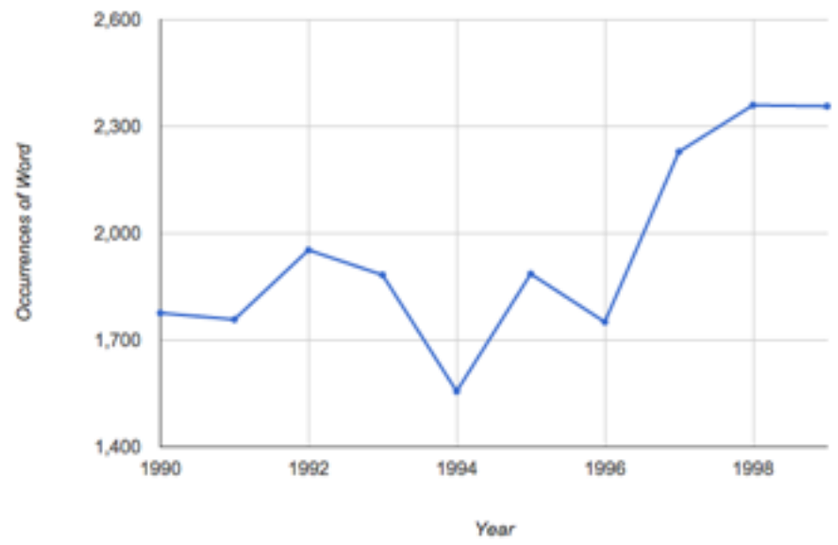
Completed analysis tasks

You have no analysis tasks already completed.

Dataset: Nature '90s

Occurrences of evolution, plotted by year

Download in CSV format



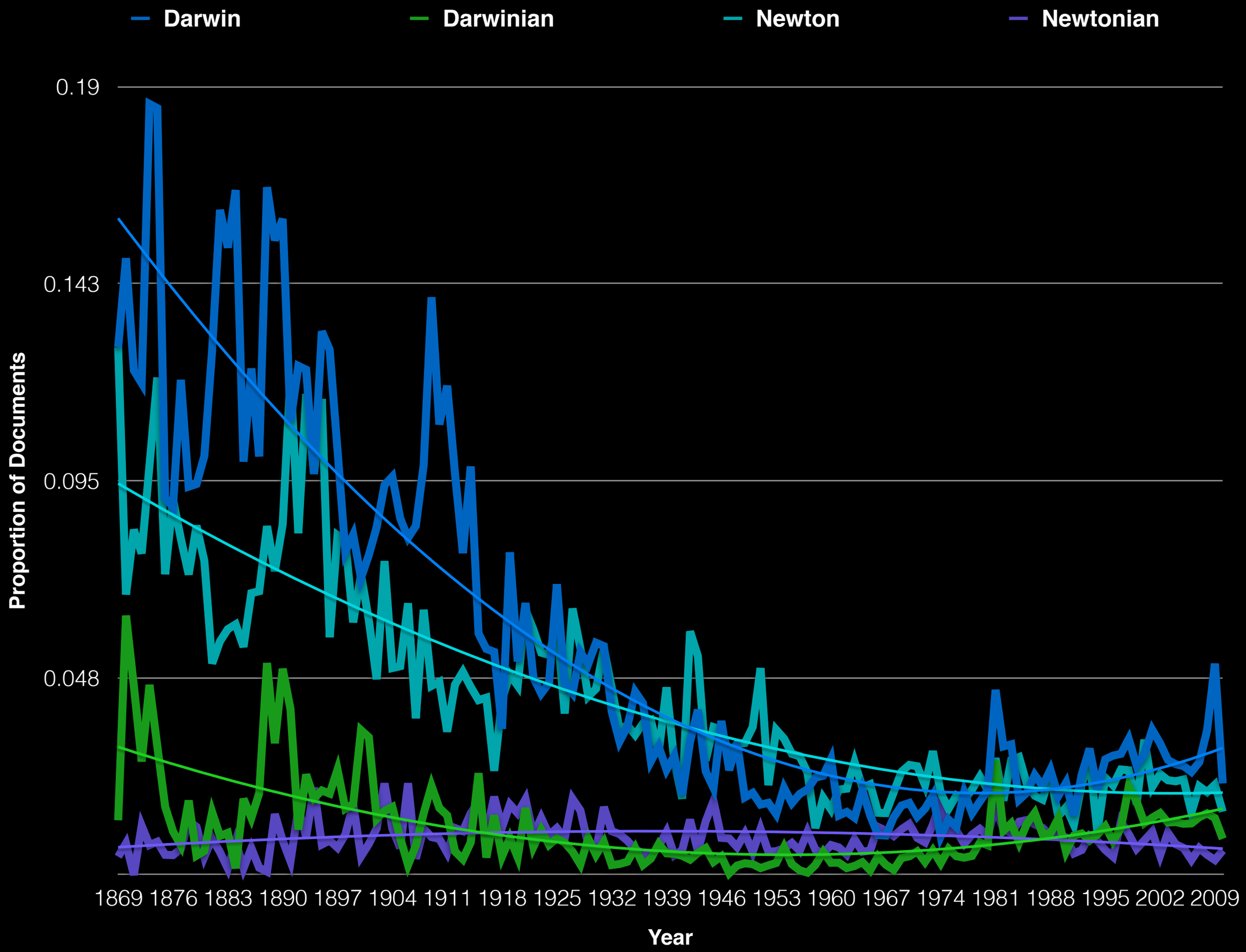
Year	Occurrences of Word
1990	1776
1991	1758
1992	1953
1993	1883
1994	1556
1995	1886
1996	1751
1997	2229
1998	2360
1999	2357

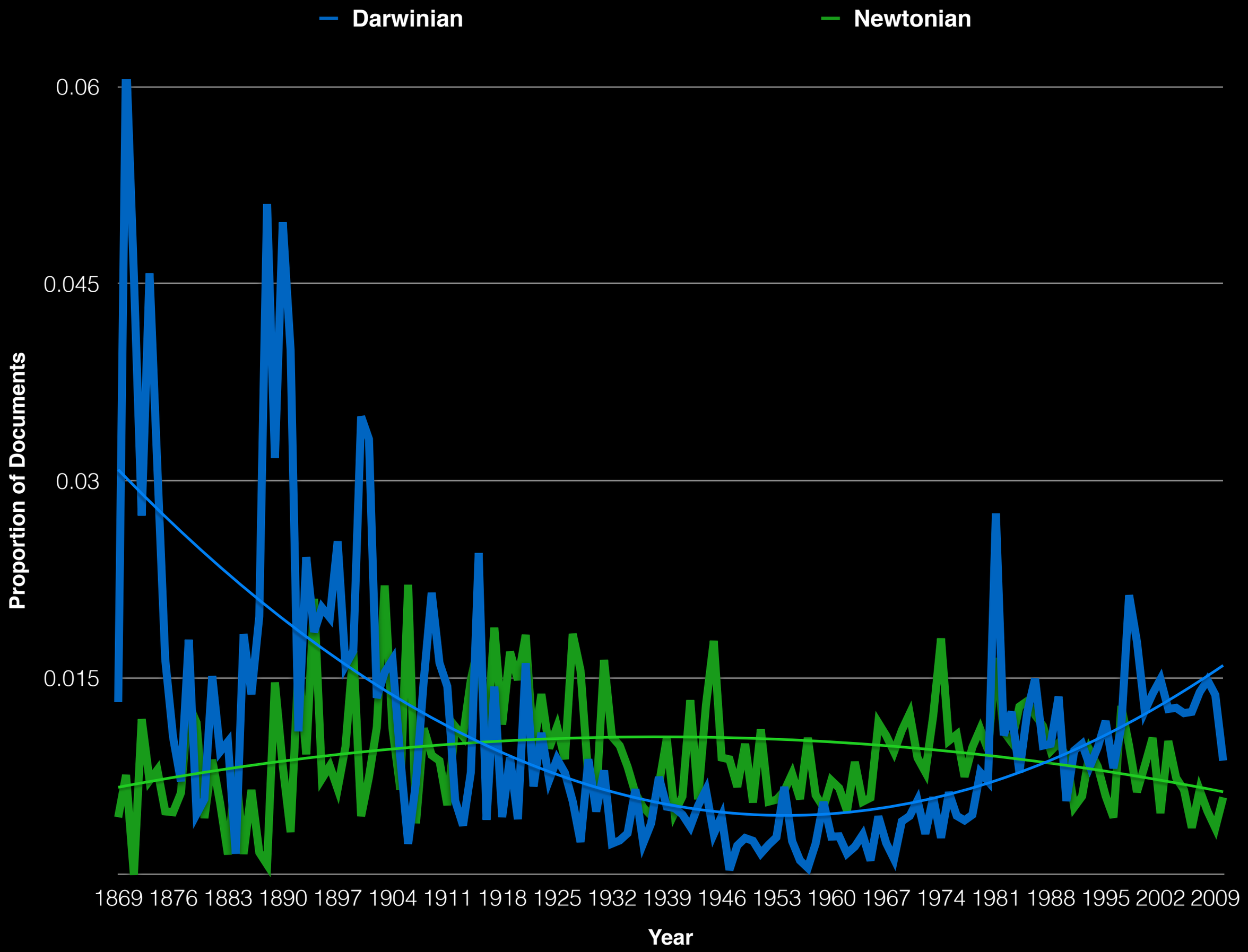
4. Preliminary results of evoText

Darwinism must die so that evolution may live

Carl Safina, 2009

“evolution can seem uniquely stuck on its founder. We don’t call astronomy Copernicism, nor gravity Newtonism”


Darwin
Darwinian
Newton
Newtonian
0.19
0.143
0.095
0.048
Proportion of Documents
1869 1876 1883 1890 1897 1904 1911 1918 1925 1932 1939 1946 1953 1960 1967 1974 1981 1988 1995 2002 2009
Year


Darwinian
Newtonian
0.06
0.045
0.03
0.015
Proportion of Documents
1869 1876 1883 1890 1897 1904 1911 1918 1925 1932 1939 1946 1953 1960 1967 1974 1981 1988 1995 2002 2009
Year

Term Frequencies

1869-1878	1935-1944	2000-2009
darwinian theory	darwinian theory	darwinian evolution
darwinian hypothesis	darwinian conception	darwinian selection
darwinian principles	darwinian theories	darwinian theory
darwinian theories	darwinian books	darwinian evolutionary
darwinian difficulties	darwinian evolution	darwinian fitness
darwinian toy	darwinian sense	darwinian approach
darwinian theorist	darwinian causes	darwinian world
darwinian views	darwinian days	darwinian medicine
darwinian doctrines	darwinian zoologists	darwinian struggle
darwinian argument	darwinian relics	darwinian look
newtonian theory	newtonian mechanics	newtonian physics
newtonian system	newtonian law	newtonian mechanics
newtonian law	newtonian liquids	newtonian gravity
newtonian colour	newtonian focus	newtonian fluid
newtonian equatorial	newtonian attraction	newtonian rheology
newtonian method	newtonian theory	newtonian fluids
newtonian basis	newtonian dynamics	newtonian dynamics
newtonian writers	newtonian system	newtonian viscosity
newtonian doctrine	newtonian era	newtonian models
newtonian telescope	newtonian acceleration	newtonian time

Term Frequencies

1869-1878	1935-1944	2000-2009
darwinian theory	darwinian theory	darwinian evolution
darwinian hypothesis	darwinian conception	darwinian selection
darwinian principles	darwinian theories	darwinian theory
darwinian theories	darwinian books	darwinian evolutionary
darwinian difficulties	darwinian evolution	darwinian fitness
darwinian toy	darwinian sense	darwinian approach
darwinian theorist	darwinian causes	darwinian world
darwinian views	darwinian days	darwinian medicine
darwinian doctrines	darwinian zoologists	darwinian struggle
darwinian argument	darwinian relics	darwinian look
newtonian theory	newtonian mechanics	newtonian physics
newtonian system	newtonian law	newtonian mechanics
newtonian law	newtonian liquids	newtonian gravity
newtonian colour	newtonian focus	newtonian fluid
newtonian equatorial	newtonian attraction	newtonian rheology
newtonian method	newtonian theory	newtonian fluids
newtonian basis	newtonian dynamics	newtonian dynamics
newtonian writers	newtonian system	newtonian viscosity
newtonian doctrine	newtonian era	newtonian models
newtonian telescope	newtonian acceleration	newtonian time

How do concepts and practices become established in the journal literature?

How do concepts and practices become established in the journal literature?

Taxon

Level of organization

Process

Theory

General

How, why, and where do new theoretical terms enter into the scientific vocabulary?

Test case: Fitness

Hypothesis:

Fitness as a central theoretical term should:

1. Enter the literature sometime after the synthesis

2. Become established in evolution-focused journals first (especially highly theoretical journals) then enter into journals that are more general, empirical, or taxon-focused

American Naturalist

Behavior

Ecology

Evolution

Journal of Mammalogy

Nature

Theory Focused

American Naturalist

Behavior

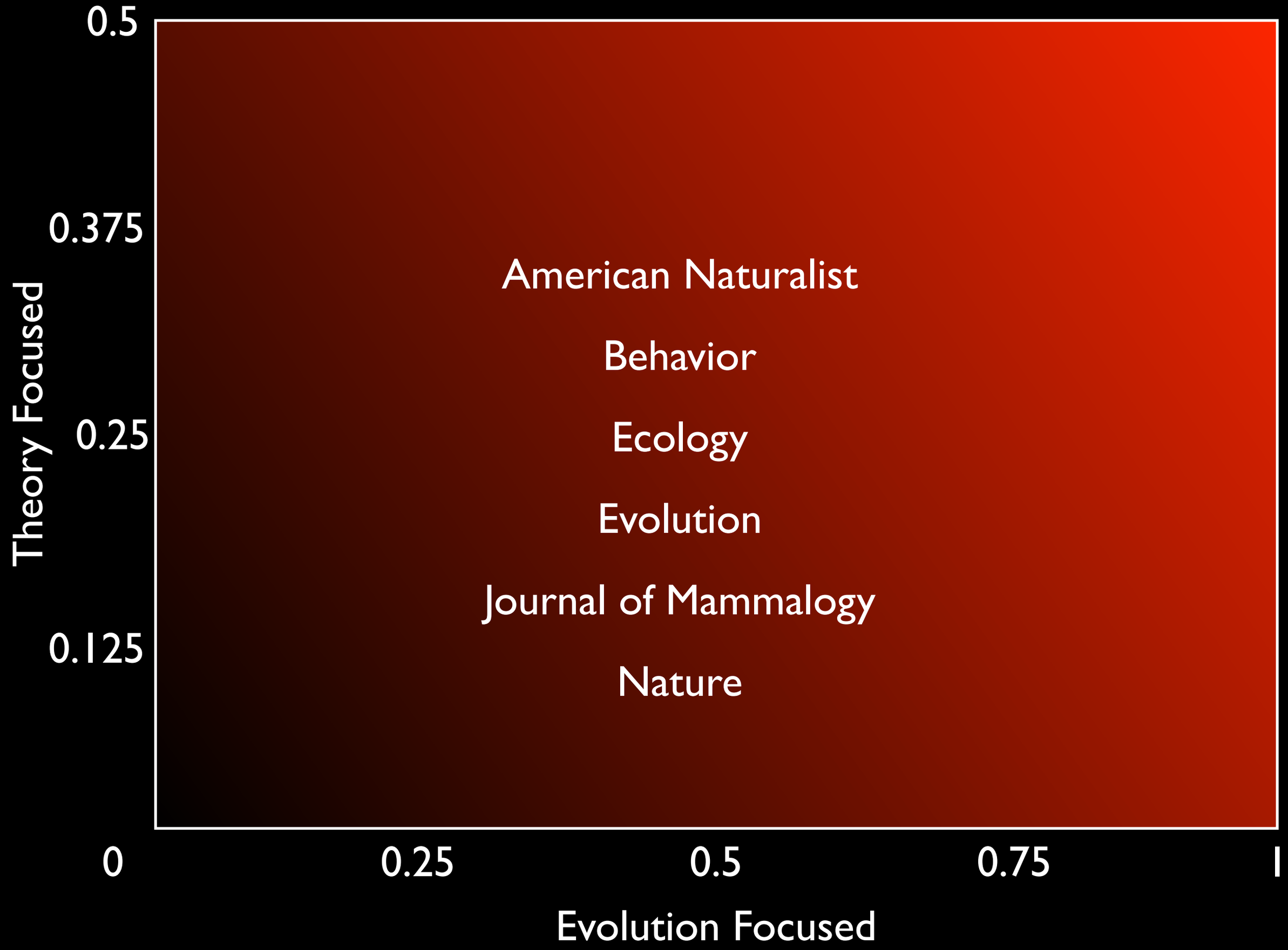
Ecology

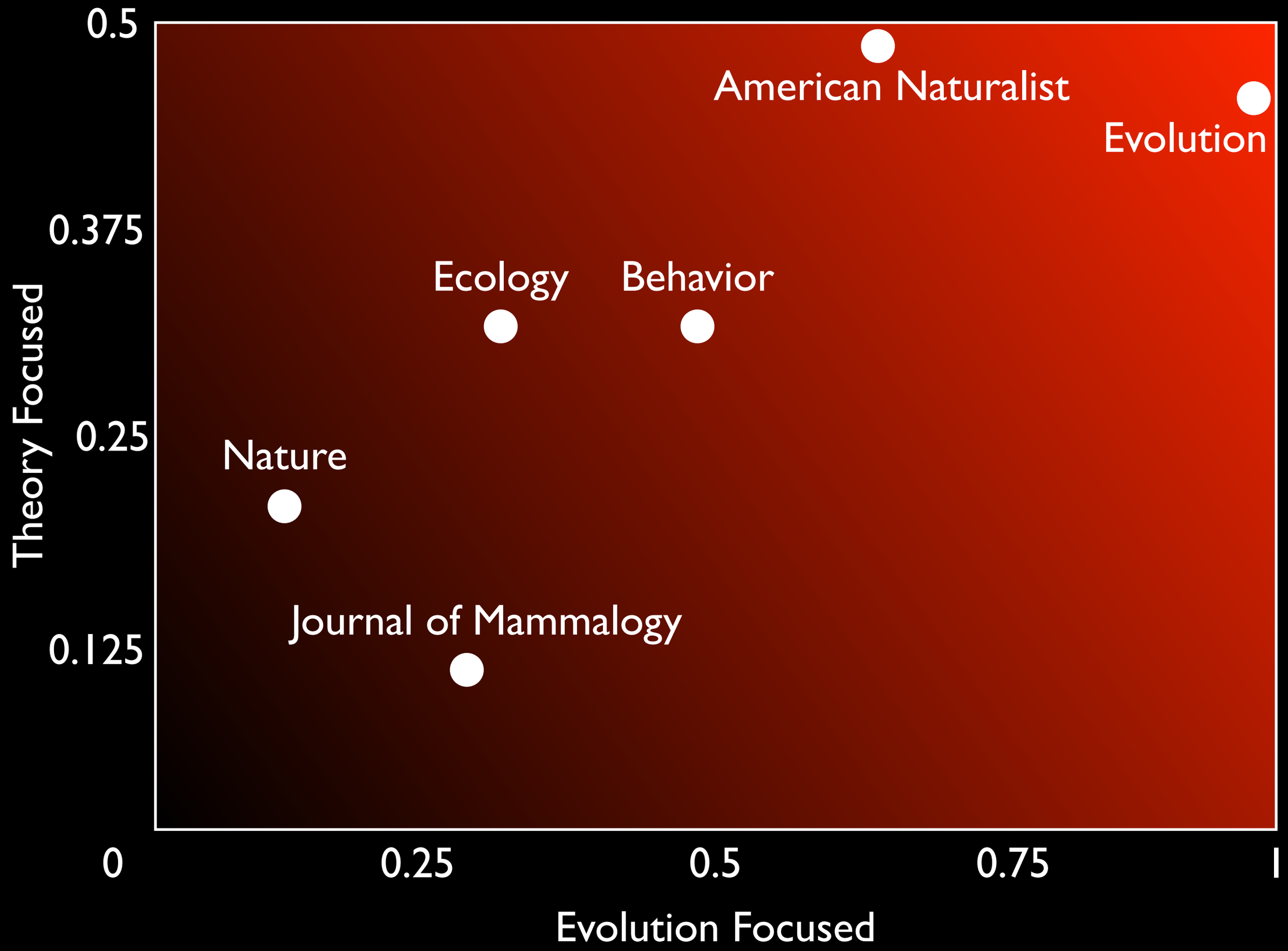
Evolution

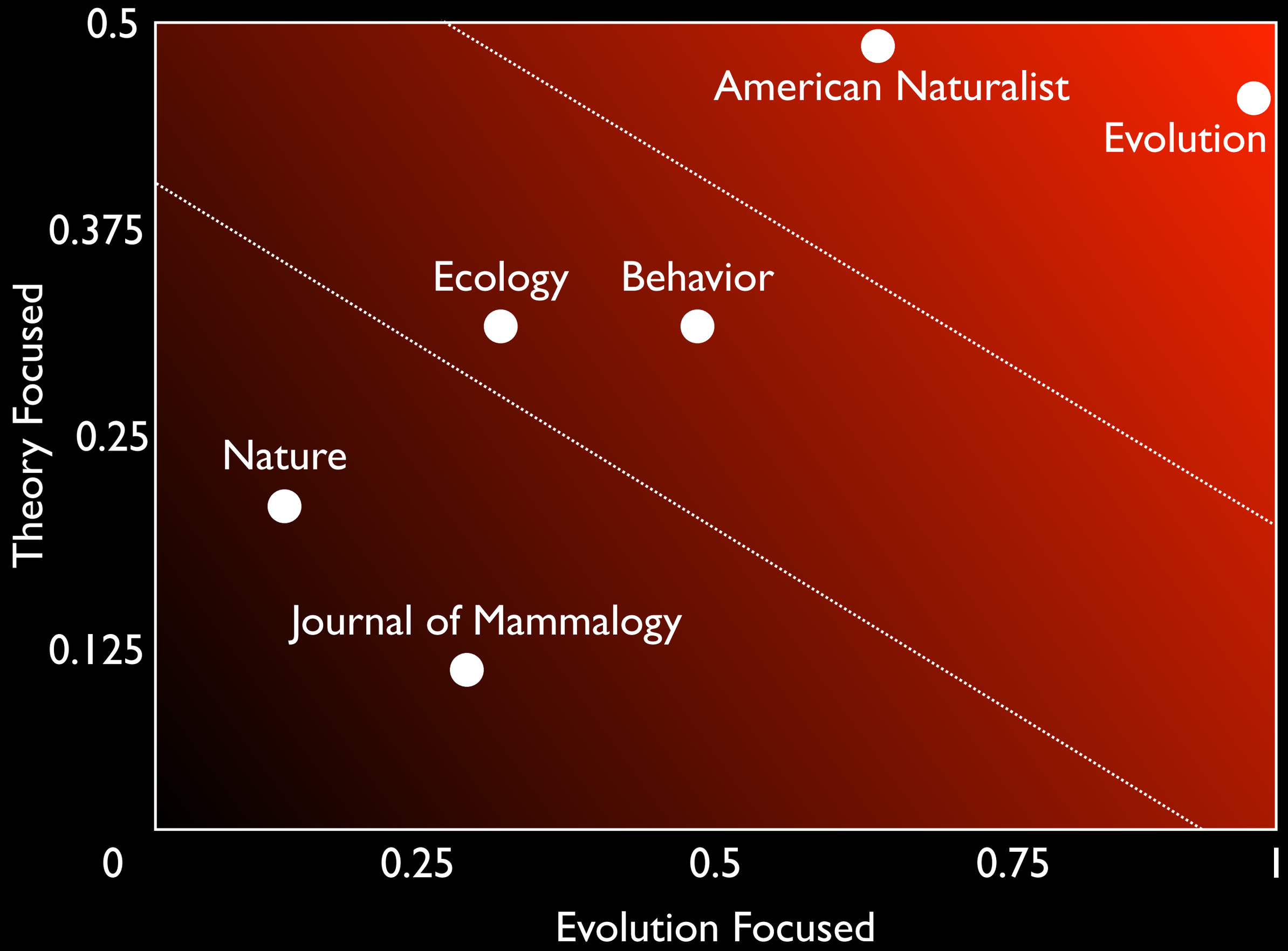
Journal of Mammalogy

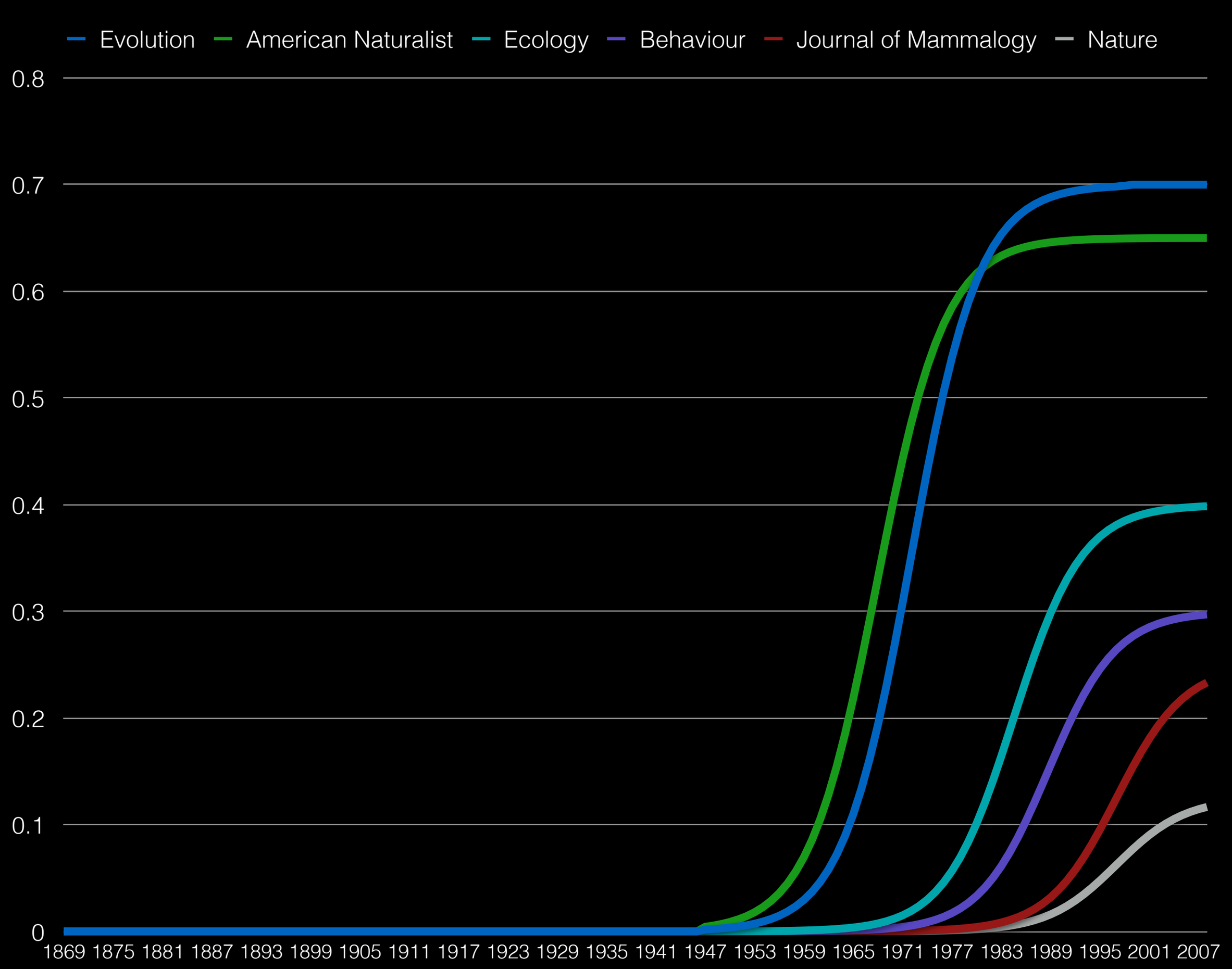
Nature

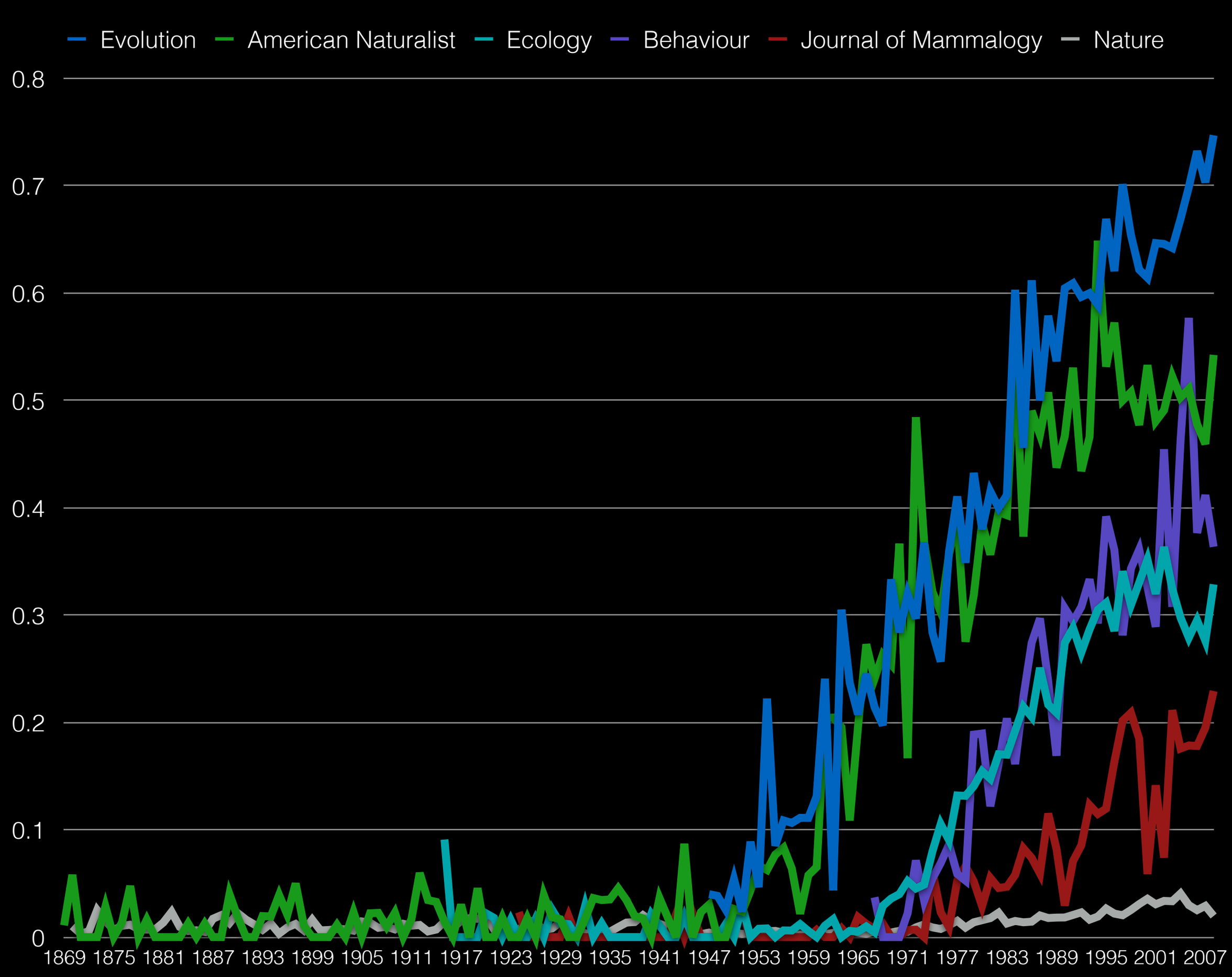
Evolution Focused

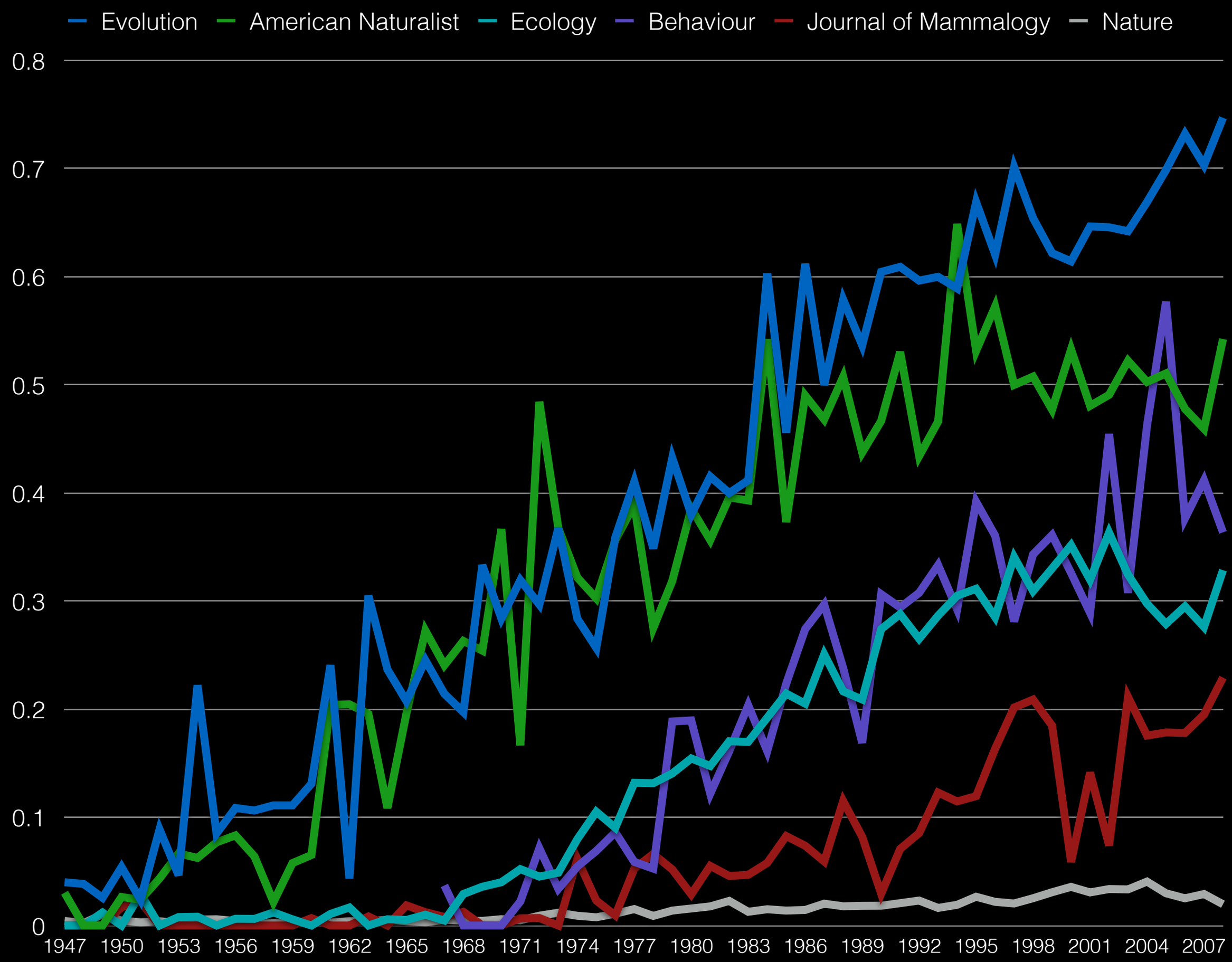

0.5
0.375
0.25
0.125
0
Theory Focused
American Naturalist
Behavior
Ecology
Evolution
Journal of Mammalogy
Nature
0.25
0.5
0.75
1
Evolution Focused

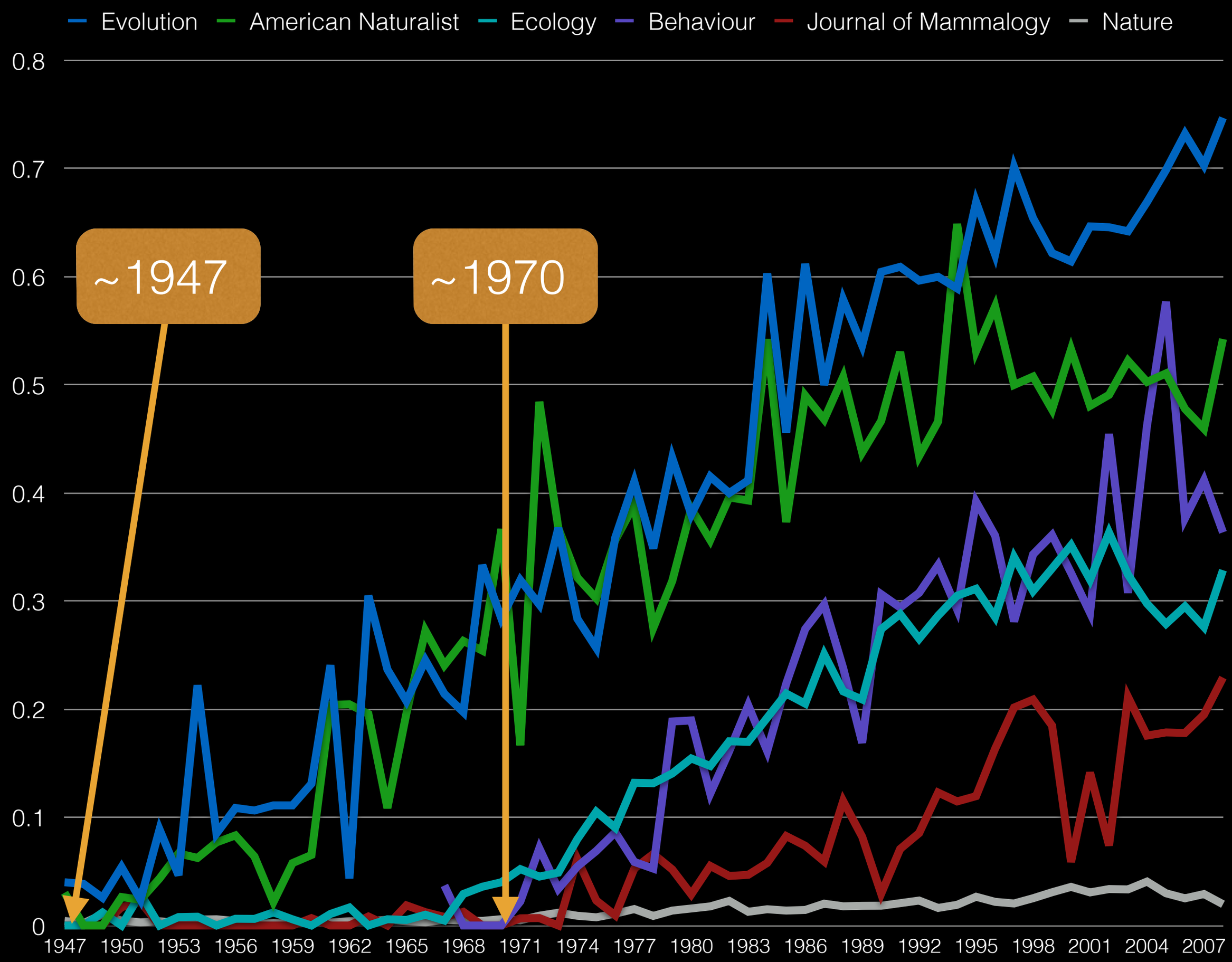

0.5
0.375
0.25
0.125
0
0.25
0.5
0.75
1
Theory Focused
Evolution Focused
American Naturalist
Evolution
Ecology
Behavior
Nature
Journal of Mammalogy


0.5
0.375
0.25
0.125
Theory Focused
American Naturalist
Evolution
Ecology
Behavior
Nature
Journal of Mammalogy
0
0.25
0.5
0.75
1
Evolution Focused


Evolution
American Naturalist
Ecology
Behaviour
Journal of Mammalogy
Nature
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1869 1875 1881 1887 1893 1899 1905 1911 1917 1923 1929 1935 1941 1947 1953 1959 1965 1971 1977 1983 1989 1995 2001 2007


Evolution
American Naturalist
Ecology
Behaviour
Journal of Mammalogy
Nature
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1869 1875 1881 1887 1893 1899 1905 1911 1917 1923 1929 1935 1941 1947 1953 1959 1965 1971 1977 1983 1989 1995 2001 2007


Evolution
American Naturalist
Ecology
Behaviour
Journal of Mammalogy
Nature
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1947
1950
1953
1956
1959
1962
1965
1968
1971
1974
1977
1980
1983
1986
1989
1992
1995
1998
2001
2004
2007


Evolution
American Naturalist
Ecology
Behaviour
Journal of Mammalogy
Nature
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
~1947
~1970
1947 1950 1953 1956 1959 1962 1965 1968 1971 1974 1977 1980 1983 1986 1989 1992 1995 1998 2001 2004 2007

Where does novelty in science come from?

Are particular kinds of journals the chief sources of novelty?

How revolutionary is science?

Does change come primarily from young or established researchers?

Do journals like *Nature, Science*, and *Cell* help or hinder science?

Nobel winner declares boycott of top science journals

Randy Schekman says his lab will no longer send papers to Nature, Cell and Science as they distort scientific process

Ian Sample, science correspondent

Follow @iansample Follow @guardian

The Guardian, Monday 9 December 2013 14.42 EST

Jump to comments (299)

Randy Schekman, centre, at a Nobel prize ceremony in Stockholm. Photograph: Rob Schoenbaum/Zuma Press/Corbis

Schekman said pressure to publish in “luxury” journals encouraged researchers to cut corners and pursue trendy fields of science instead of doing more important work. The problem was exacerbated, he said, by editors who were not active scientists but professionals who favoured studies that were likely to make a splash.

Nobel winner declares boycott of top science journals

Randy Schekman says his lab will no longer send papers to Nature, Cell and Science as they distort scientific process

Ian Sample, science correspondent

Follow @iansample Follow @guardian

The Guardian, Monday 9 December 2013 14.42 EST

Jump to comments (299)

Randy Schekman, centre, at a Nobel prize ceremony in Stockholm. Photograph: Rob Schoenbaum/Zuma Press/Corbis

Schekman said pressure to publish in “luxury” journals encouraged researchers to cut corners and pursue trendy fields of science instead of doing more important work. The problem was exacerbated, he said, by editors who were not active scientists but professionals who favoured studies that were likely to make a splash.

Nobel winner declares boycott of top science journals

Randy Schekman says his lab will no longer send papers to Nature, Cell and Science as they distort scientific process

Ian Sample, science correspondent
Follow @iansample Follow @guardian

The Guardian, Monday 9 December 2013 14.42 EST

Jump to comments (299)

Randy Schekman, centre, at a Nobel prize ceremony in Stockholm. Photograph: Rob Schoenbaum/Zuma Press/Corbis

Schekman said pressure to publish in “luxury” journals encouraged researchers to cut corners and pursue trendy fields of science instead of doing more important work. The problem was exacerbated, he said, by editors who were not active scientists but professionals who favoured studies that were likely to make a splash.

How could one operationalize splashiness or trendiness?

What effect does splashiness of trendiness have on science?

This is just the beginning…

What are your ideas for evoText?

Thank You