



Ecological Dynamics of Discourse in Scientific Communities: Co-evolution of Conceptual and Social Networks

Drew B. Margolin

Annenberg School of Communication

University of Southern California



Overview



- Research Motivation
- Theoretical Argument
- Methodological Proposal
- Preliminary Results



Motivation



Research Question:

What are the structural signatures of impending growth or decline in paper production on a particular scientific topic?



With emphasis...



What are the *structural* signatures of impending *growth* or *decline* in *paper production* on a particular topic?

Three important aspects:

1. Signatures are structural rather than content based
 - The assumption of scientometrics (de Solla Price, 1965; Börner, Boyack, & Chen, 2003)
2. Substantial interest in growth of scientific fields
 - But little emphasis on abandonment or decline (Bettencourt et al., 2009)
3. Dependent variable is paper production (Latour & Woolgar, 1978)
 - Temporarily accepted “knowledge”




Structural indicators of scientific growth




Kuhn (1962) provided first semi-structural theory:

- It is neither necessary nor sufficient for knowledge claims to match a “real world”
- Rather...science is produced by relations amongst the extant concepts and propositions in a field (Feyerabend, 1977)
 1. To grow, a field requires a sufficient number of *paradigmatic premises* to foster and maintain collaboration
 2. To survive, a field must stave off the accumulation of *anomalies*
- But Kuhn’s relativistic approach leaves no room for the “third-party observer” (Maturana & Varela, 1980; Luhmann, 2002)
 - Paradigmatic premises and anomalies are endogenously defined by the community of scientists
 - ...similar to the problem of defining “ontologies” (van Atteveldt, 2008)



Recent findings suggests we may have signatures of these phenomena



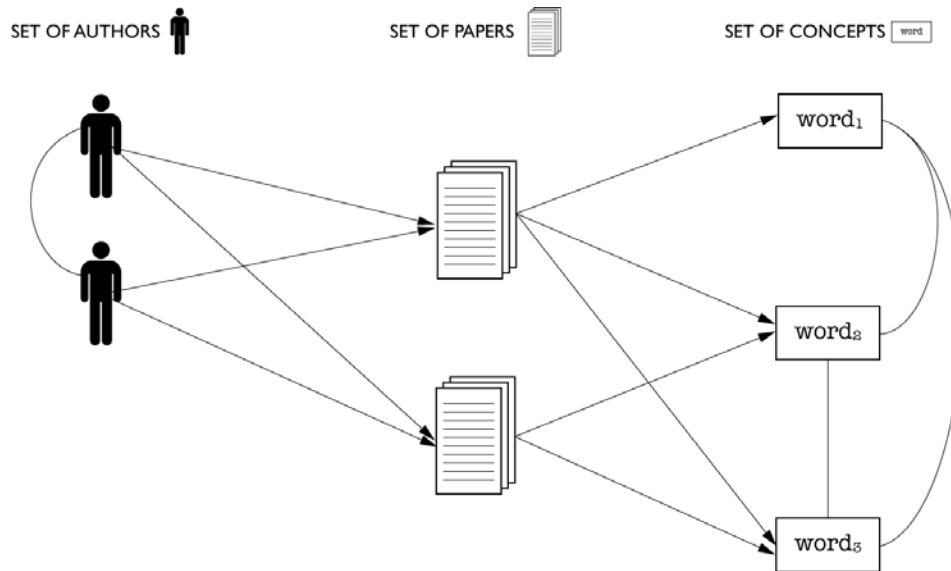
- Paradigmatic premises
 - Concepts with large authorial extent (Roth & Bourguine, 2006)
 - Concepts in which all authors are “expert” (Tamarasco, Roth, and Cointet, 2010)
- Anomalies
 - New fields turnover new words faster than established fields (Bonaccorsi, 2008)
 - What cannot be found must be invented
 - Invention of new concepts tend to have names that carry no a priori meaning/association (e.g. “name” of the inventor) (Frigotto & Ricabonni, 2010)
- But we need more theory to refine distinctions...
 - Between paradigmatic concepts and meaningless concepts
 - Between breakthroughs and anomalies

Broad theoretical proposal...

- Consider scientific production as a co-evolutionary process proceeding by variation, selection, and retention (Darwin, 2002 [1859]; Campbell, 1965; Hazen & Eldridge, 2010)

Scientists produce papers

A portion of these papers are *selected* by reviewers for publication





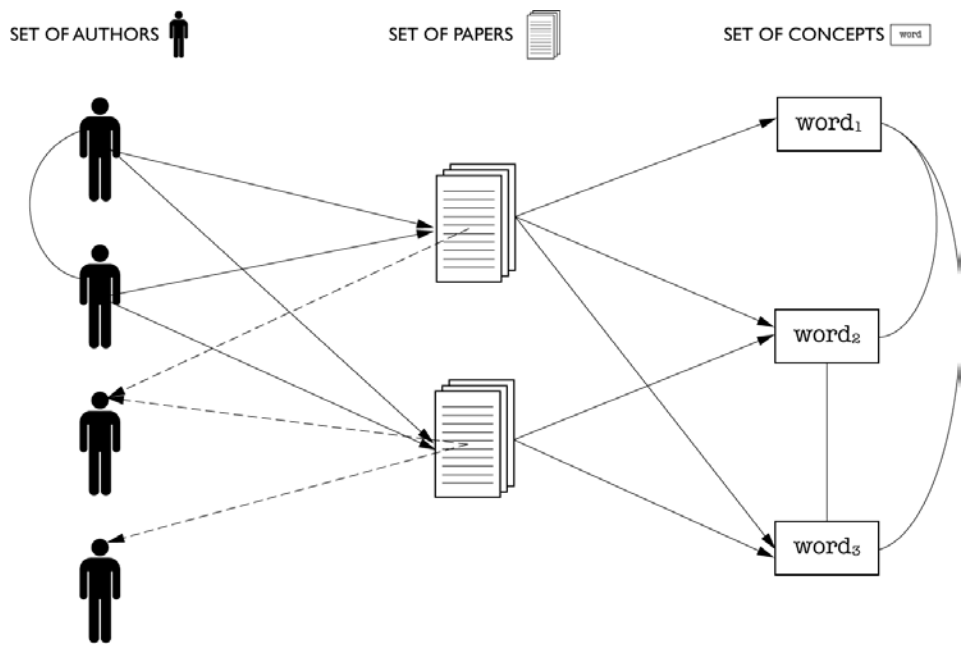
Broad theoretical proposal...



Scientists produce papers

A portion of these papers are *selected* by reviewers for publication

Other scientists *search* the selected, published papers for *novel connections* between ideas... (Swanson, 1986; Kostoff, 2008)

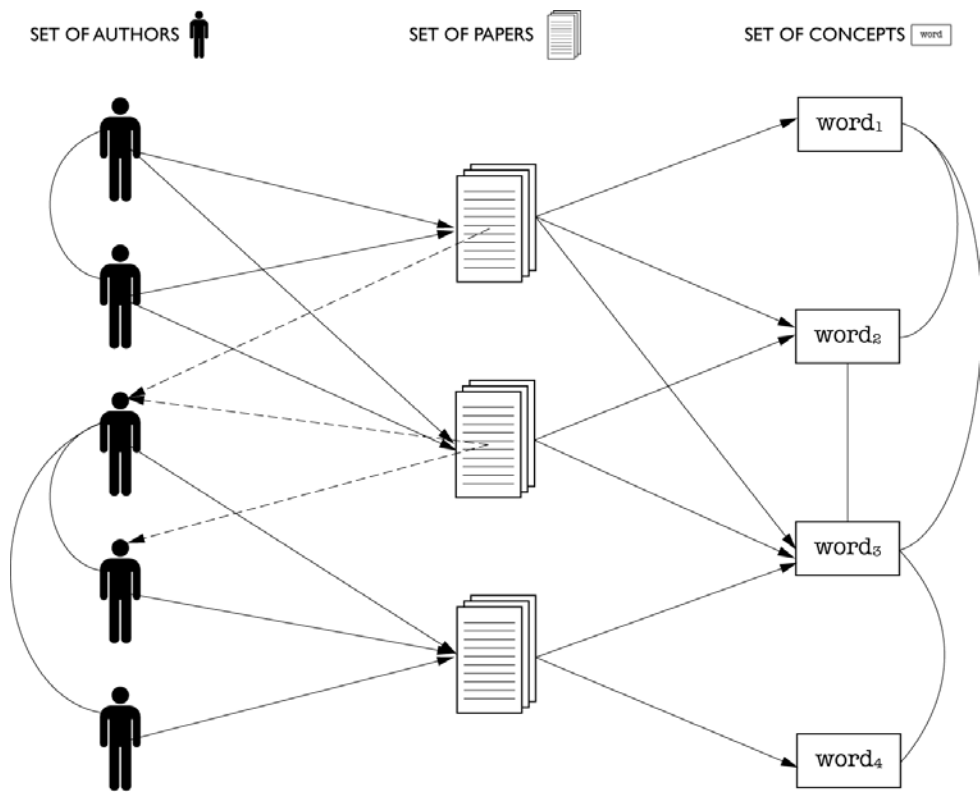




Broad theoretical proposal...



- Scientists produce papers
- A portion of these papers are *selected* by reviewers for publication
- Other scientists *search* the selected, published papers for *novel connections* between ideas... (Swanson, 1986; Kostoff, 2008)
- ... and produce new papers





Evolutionary and ecological models have been successful in other fields



- Organizational ecology and neo-institutional theory (DiMaggio & Powell, 1983; Hannan & Freeman, 1977; Meyer & Rowan, 1977)
 - Organizational communities transform via replacement rather than adaptation
 - Organizations use inter-organizational linkages to supplement their inability to adapt (Powell et al., 2005; Monge, Heiss, & Margolin, 2008)
 - Organizations change through periods of convergence and re-orientation (Miller & Friesen, 1981; Tushman & Romanelli, 1985)



Apply Evolutionary Mechanisms to Existing Structure of Research on a Topic



- Scope of Variation
 - Individuals search for ideas using established conceptual connections as the starting point (McKoon & Ratliff, 1992; 1998; Lowe & McDonald, 2000).
 - conceptual network of a text suggests the attention required to search it for meaning (Corman et al., 2002)
 - Thus, a conceptual network can be evaluated for the scope of probable “recombinations” of concepts and propositions built from links between concepts in the **conceptual network** (Hazen & Eldrige, 2010)
- Strength of Selection
 - Concepts are “rules” or norms for what can be said
 - **Social network** determines the *acceptability* of novel combinations of concepts by governing the permissiveness of the community (Coleman, 1988; Burt, 2005)



Leads to...



- The Shape of Retention

- Combinations of concepts that are repeatedly selected across research on the topic “accumulate” into the new, **conceptual network**
- The organization of these networks can be measured and compared across fields (Leydesdorff & Hellsten, 2005)
- This organization is a shared good or predicament



The Evolutionary Frame..



- The retained conceptual network is knowledge to students, technicians, ...etc.
 - ...but it is only a “starting point” for researchers
- The current conceptual network “always fits” the available data (Feyerabend, 1976)
 - ... but does it do so in a productive way?
 - At what cost was the fit obtained?
 - Are recombinations easier or more difficult?
 - How accessible are other areas of related knowledge?
 - How many individuals have mastery over the requisite concepts?



Example... conceptual possibilities for an unexpected finding...



Finding:

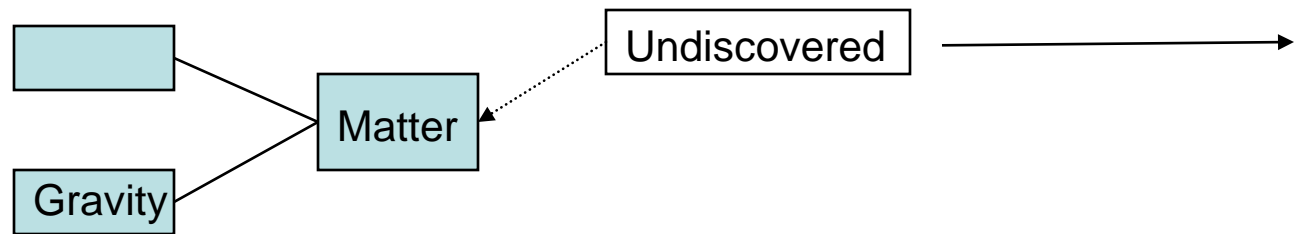
- The gravitational force at work in the universe is far greater than can be provided by identifiable matter.

Possible conceptual “compensations” to account for the finding

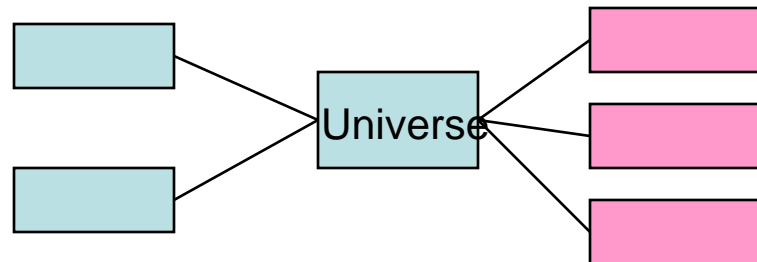
1. Hidden, unseen “dark matter”
2. “Multiple universes”
3. “Gravity” is actually two forces, one generated by “mass,” another generated by a “sense-non”

Each semantic compensation has a distinct consequence for the conceptual network

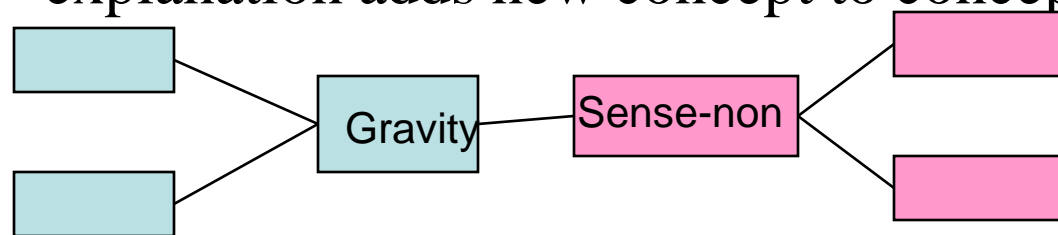
- “Dark matter” preserves standard matter-gravity relationship but promises new observations...




- “Multiverse” explanation loosens concept of “universe”




- “Sense-non” explanation adds new concept to concept of “gravity”





The accumulation of these choices has consequences



Depending on the existing configuration of the conceptual network and the discrimination provided by the social network, it is possible to have...

- A conceptual network that is too disconnected to be searchable (requiring long jump evolution) --> Low Paper Production
- A conceptual network that is too connected to be informative (noise is selected for) --> Low Paper Production
- A conceptual network that is “just right” --> High paper production



The consequences of accumulated retention

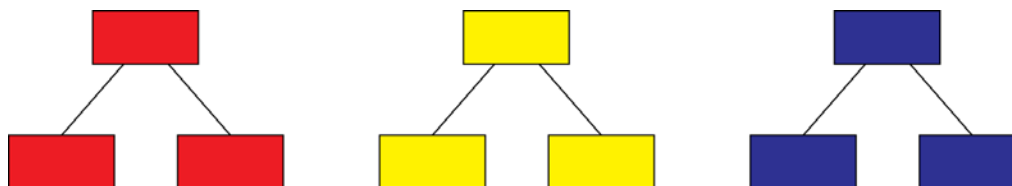


- How searchable is the conceptual network, how easily can novel combinations be found?

Signatures of poor search-ability...

- Disconnectedness of network (no recognized association between core concepts)

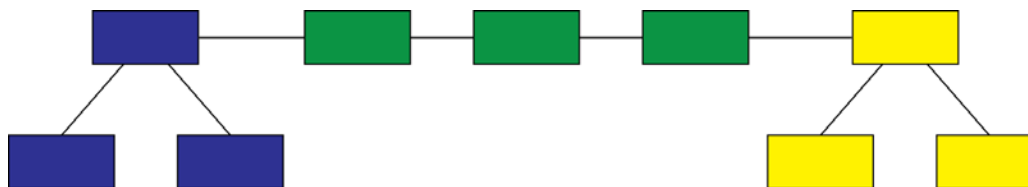
Lost Syllogisms (Swanson, 1986)



- High between-ness centrality (all claims must use same core concepts)

“Frozen Components” (Kauffman, 1993)

Hegemonic Concepts (Callon et al., 1983)





What disciplines this growth?



The social network amongst authors and reviewers.

- Conceptual usage is a rule-based phenomenon (Searle, 1969)
- Cohesiveness of social networks are the source of rule conformity (Coleman, 1988; Burt, 2005)

That is, the community of scientists establish and maintain their standing in the community based on what their judgments of what conceptual combinations can be allowed.

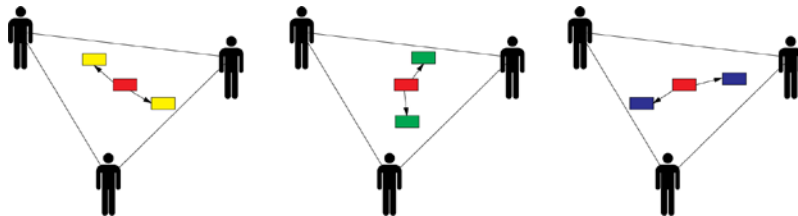
How discriminating is the social network?

- How useful is the existing set of findings for discovering plausible, meaningful new knowledge claims?
- Weak selection” can trap populations at the periphery of fitness peaks (Kauffman, 1993)

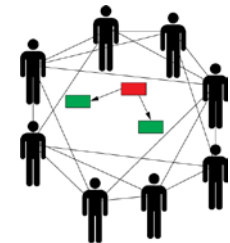
Signatures of weak Selection:

- Disconnectedness of social network

No enforcement of norms (Coleman, 1988; Burt, 2005; Uzzi, 1997)

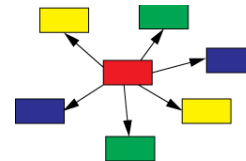


-versus-



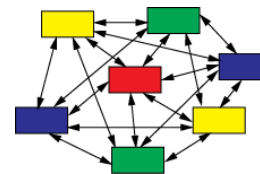
- High degree in conceptual network

Meanings are flexible



- High entropy in the conceptual network (Tutzauer, 2007)

No concept is a guide to any other concept





Preliminary Analysis



- Data
 - Papers gathered from WoS w/keyword “fullerenes” (1991 - 2006)
 - conceptual network gathered from co-occurrence in abstracts (every 3 years)
 - Social network of links between authors
- Pre-processing in Automap 2.7.2 (Carley, Diesner & DeReno, 2006).
 - Discard Salton and Buckley stop words
 - No stemming
 - Any co-occurrence in an abstract counts as a link (dichotomous network)

Descriptive Statistics

Conceptual Network

	1991	1994	1997	2000	2003	2006
Articles	44	399	495	488	491	578
Abstracts Available	41	369	462	458	464	547
Tokens	2,217	22,534	28,448	30,102	32,984	39,964
Concepts Used	1095	5694	6857	7225	7339	8595
Concepts Repeated 2x or more	382	2586	3158	3285	3496	4001

Social Network

	91, '94	91, '94, '00	94, '97, '00	97, '00, '03	00, '03, '06
Unique Authors	1118	2299	3260	3333	3756
Density	0.41%	0.19%	0.14%	0.16%	0.15%
Closed Triangles	28,680	50,508	74,640	94,056	122,746
Components	279	574	718	605	636
Largest Component	151	151	681	1249	1704
% in Largest	14%	7%	21%	37%	45%
Fragmentation	0.974	0.991	0.954	0.858	0.793



Preliminary Patterns



	1991	1994	1997	2000	2003	2006
Retained Concepts	382	2586	3158	3285	3496	4001
Degree Density	0.125	0.049	0.043	0.043	0.043	0.041
Avg. Degree per node	95	254	270	282	301	329
Normalized Avg. Degree per node	113	266	281	274	271	288
Degree centralization	0.65	0.64	0.63	0.68	0.60	0.59
Between-ness centralization	0.13	0.08	0.08	0.06	0.06	0.05
Close-ness centralization	0.84	0.68	0.65	0.57	0.58	0.62
# of unreachable pairs	381	15,499	34,703	42,648	31,436	59,958
Avg path distance (reachable pairs)	1.91	2.00	2.02	2.03	2.02	2.01
Diameter	3	4	5	5	4	4

Conceptual networks shows...

- Constant density
- Declining between-ness and closeness centralization
- Increasing, then decreasing diameter

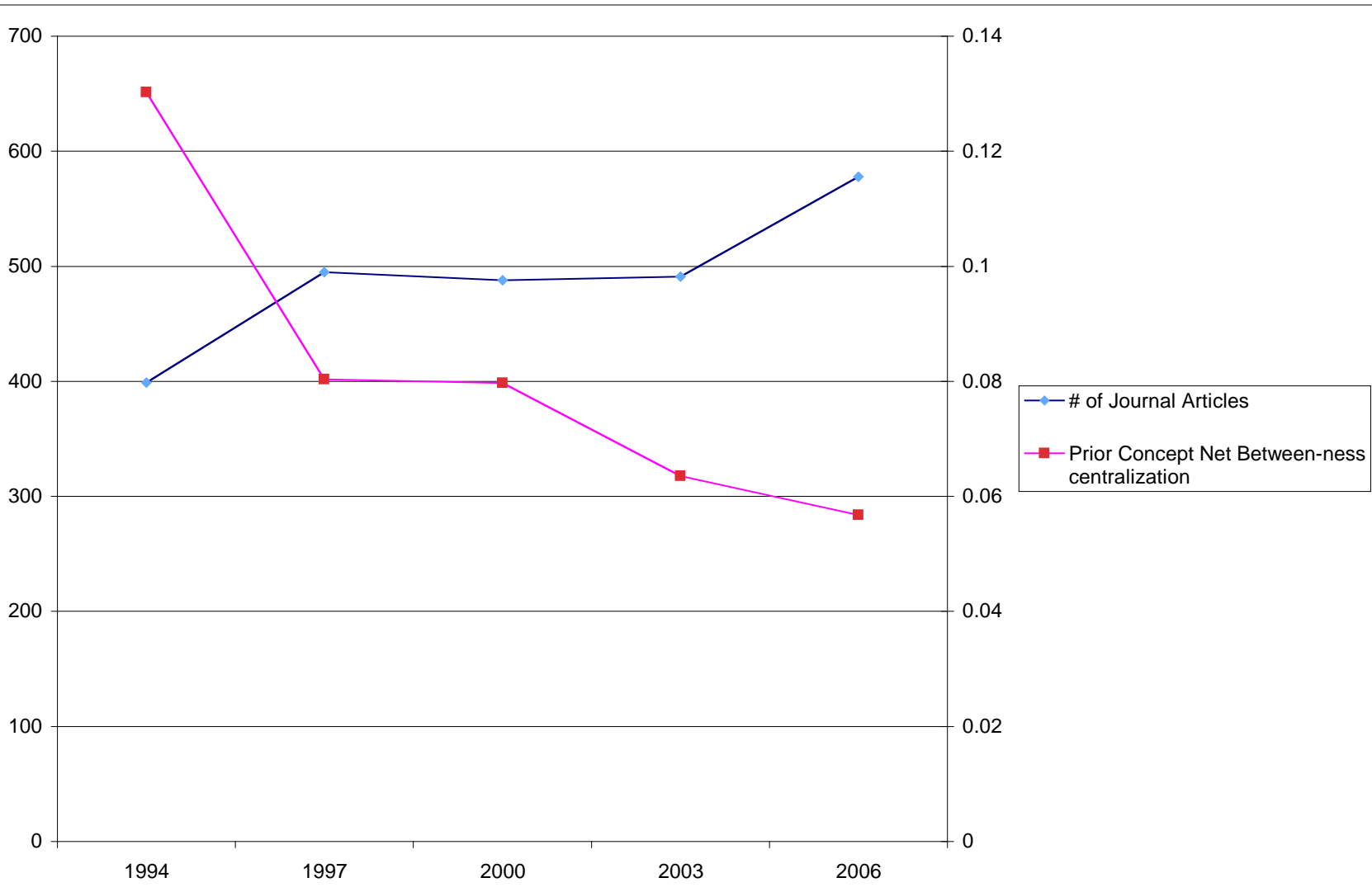
As invisible college forms, concepts are repeated



Use of repeated concepts is “disciplined” by social connectedness



Reduced between-ness in semantic network fosters paper production





Thank you for your attention!



Please e-mail me at dmargoli@usc.edu with questions or comments!

Acknowledgements: Thanks to Peter Monge, Janet Fulk, and David Kempe for their guidance on this project!