ERC Advanced Grant 2018
Extended synopsis of the scientific proposal [Part B1a]

An algorithmic framework for reducing bias and polarization in online media
(REBOUND)

Aristides Gionis

Section a. Background and objectives

Social media play a critical role in today's information society, not only by connecting people with their friends, but also by providing a medium where information is disseminated and public opinion is shaped. Initially it seemed that giving ordinary citizens the means to create content of their own and share their opinion publicly can have only positive effects: increase the exposure to diverse ideas and improve the democratic process. However, during the past few years we have witnessed that the rise of online media has led to a series of undesirable phenomena, such as creation of information silos and increased polarization. More and more digital citizens find themselves in online environments where they only see opinions that confirm their preexisting beliefs, and as a result, they become ideologically segregated.

These negative aspects of social media have drawn a lot of attention recently. There has been a considerable amount of criticism, skepticism, discussion in public forums, as well as calls to fix the problems. Prominent media outlets, researchers, and think tanks have contemplated whether "*social media is a threat to democracy*."[1] Given all these concerns a natural question arises.

How can we address the deficiencies encountered in today's online platforms and how can we create environments in which online media enable exchange of alternative views and promote constructive deliberation?

This is certainly a complex problem that requires parallel work and cooperation of multiple parties: owners of social-media platforms and traditional media outlets, journalists, policy makers, educators, as well as independent academics. It is also a cross-disciplinary research challenge that requires expertise in sociology, political science, economics, engineering, design, and *computer science*.

The REBOUND project aspires to contribute in addressing this complex problem by designing methods to perform large-scale data analysis so as to shed light to phenomena such as bias, polarization, conflict, and information silos in online media. It will also contribute with methods that give users easier access to opposing ideas and incentives to explore and understand alternative viewpoints. More concretely, the project has the following goal.

High-level goal of REBOUND**:** We will develop theoretical foundations and a concrete set of algorithmic techniques to address deficiencies in today's online media. We will develop methods to discover structure and patterns of segregation, conflict, and closeness in social-media systems. We will address the issues of reducing bias and polarization, breaking information silos, and creating awareness of users to explore alternative viewpoints.

Challenges. Achieving the project goal is an extremely challenging task for a number of reasons. First, we will be dealing with noisy data, generated by a diverse set of actors who exhibit complex behavior and may have adversarial motives. Additionally, data are highly dynamic: new events are taking place and interactions between actors are constantly evolving over time. Second, we will deal with concepts that are not well-understood and we will need to develop novel abstractions. For instance, while there is a well-motivated discussion about how to combat the problem of bias and polarization in online media, there is no understanding of what does this entail nor what could be the side effects of possible remedial actions. Third, we will be confronted with challenging computational tasks that lead to difficult optimization problems. Thus, it is desirable to develop methods that offer approximation guarantees or other theoretical properties.

[1] For example, see www.economist.com/leaders/2017/11/04/do-social-media-threaten-democracy, or www.weforum.org/agenda/2016/08/the-biggest-threat-to-democracy-your-social-media-feed/

Hypothesis and objectives. The project is grounded on the following hypothesis.

Hypothesis: We postulate that bias, conflict, polarization, and information silos are real phenomena in today's online media. We hypothesize that these deficiencies can be alleviated by designing and building appropriate tools, and that people can get interested to explore and engage with alternative viewpoints.

Our encompassing research objective is to test this hypothesis. To achieve this objective we aim to consolidate existing approaches, including our recent work, and push significantly the state-of-the-art in terms of models, novel problem formulations, improved algorithmic techniques, and applications. In particular, REBOUND has the following concrete objectives.

Models: Develop novel models for analyzing data on news dissemination and online discussions in online media, with the aim to obtain a deeper understanding on phenomena related to bias, polarization, and information bottlenecks. Extend standard data-mining problem formulations for this particular application domain, and formulate novel problem representations for the tasks specified in the three research thrusts of the project.

Algorithms: Develop efficient algorithmic solutions for the formulated problems. Our algorithms should take into account the rich available data representations as well as the stochastic and noisy nature of the problem domain. The proposed algorithms should be efficient, should be able to deal with uncertainty, and should offer theoretical guarantees.

Applications: Apply the developed methodology on different application scenarios and evaluate the resulting algorithms on real-world datasets. Validate proof-of-concept by showcasing findings of the methods on different use cases. Implement the developed algorithms and make them available to the scientific community.

Education: Nurture doctoral students and postdoctoral researchers. Educate them about fundamental algorithmic techniques, teach them the value of identifying important data-analysis problems, and support them becoming independent researchers.

State of the art. A number of theories have been proposed in psychology and social sciences in an attempt to *explain the mechanisms* that result in various types of bias that are present in our society. Those include the theories of *cognitive dissonance* [14], *selective exposure* [15], and *biased assimilation* [26]. All these psychological mechanisms, together with other biases, such as, *algorithmic filtering* and *personalization* [9], are connected to the phenomenon of information silos and echo chambers in online media [30]. With the wide availability of online media data, a lot of work has been devoted in studying controversy, polarization, bias, and conflict in social media. In one of the first papers, Adamic and Glance study the link structure of blog posts on the 2004 US presidential election; they provide evidence that conservative blogs are linking to each other more frequently and in a denser pattern [1]. Conover et al. study twitter data from congressional midterm elections and identify a highly-segregated partisan network structure by the means of graph partitioning [11].

Echo chambers have been shown to exist in various forms of online media such as blogs [20, 39], forums [13], and social-media sites [8, 22, 31]. An et al. analyzed the activity of twitter users who engage with political news and found that 90% of the users directly follow news media of only one political leaning [2]. In a study performed in Facebook, Bakshy et al. find that although users are exposed to a significant amount of cross-cutting content, they opt to engage with proportionately less cross-cutting content [6]. A few recent papers have addressed the problem of mitigating polarization and diffusing echo chambers. The underlying ideas involve making users aware of other users' stance on a given issue [24, 25], or revealing the bias of the news articles they read [28]. In work pioneered in my research group, we have studied the problem of reducing bias and polarization by the means of recommending links of users to follow [16], selecting users to promote campaigns [17], and maximizing diversity [4, 27].

In summary, the problem of reducing bias and polarization in online information has already received some attention in the literature. However, the area is still largely undeveloped and most works are preliminary. The topic has attracted a lot of attention in the public discussion forum, and there is exciting work ahead to be done in a range of disciplines, including with no doubt computer science.

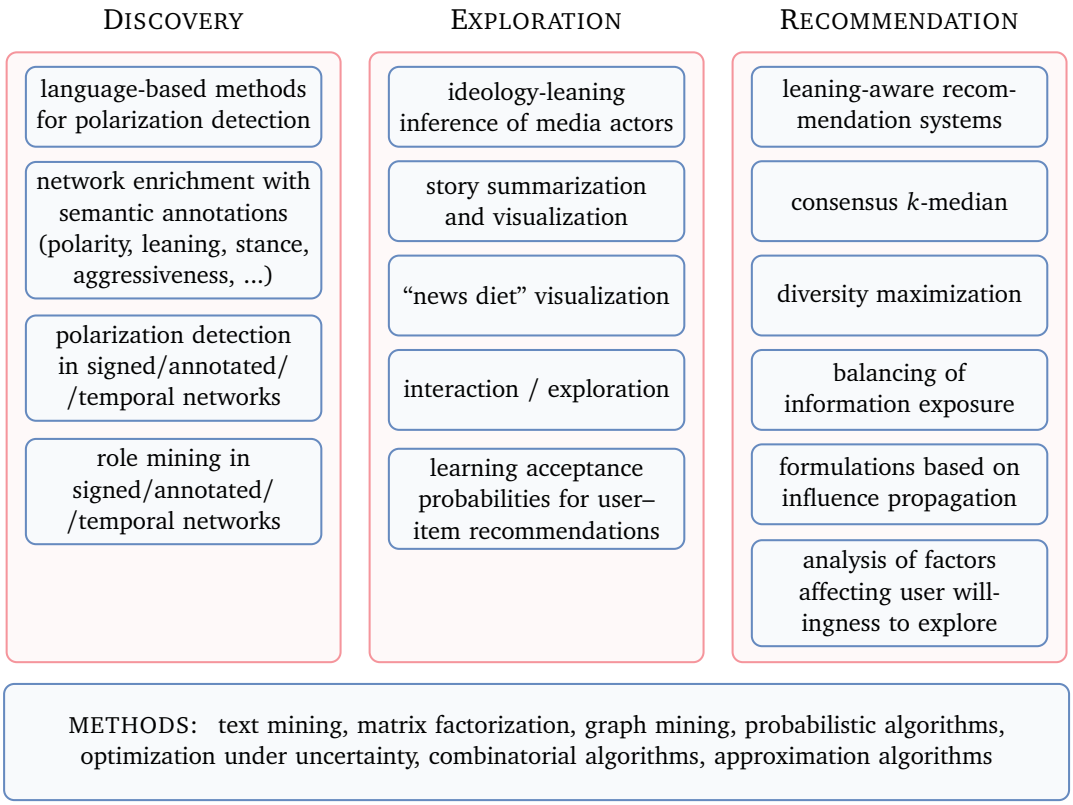

Figure 1: The structure of REBOUND, depicting the research thrusts of the project, the problem formulations we will consider, and the methods we will employ.

Section b. Methodology

The project is structured in three *research thrusts*: DISCOVERY, EXPLORATION, and RECOMMENDATION. Each thrust specifies a research domain within the project and a set of concrete problems with common motivation and common vision. While to a large extent the problems within each thrust can be studied independently, there are also strong connections and feedback from one thrust will be used to obtain higher-quality results in the other. The structure of the project is shown in Figure 1. In addition to the problems we will study within each thrust, we also show the computational methods that we will employ.

With respect to methods, emphasis will be given to combinatorial algorithms, building on the previous work of the PI on developing combinatorial methods for data-mining problems. In particular, we will consider techniques such as combinatorial optimization, optimization of submodular functions, local-search methods, greedy algorithms, dynamic programming, linear-programming and semidefinite-programming relaxations, primal-dual methods, convex optimization, etc. In addition, to deal with the stochastic nature of the data we will develop probabilistic models and algorithms that handle uncertainty. We will also focus on designing scalable methods, by applying techniques based on sampling and sketching [12]. The project also requires expertise on text analysis, natural-language processing, visualization, and user studies. These needs will be covered by hiring postdoctoral researchers with relevant expertise.

Next we overview the three research thrusts of REBOUND.

Research thrust 1: DISCOVERY

In this research thrust we will develop novel computational methods to discover patterns and structure related to bias, polarization, conflict, and information silos in online media. We will focus on mining semantically-annotated graphs with information extracted by text analysis, mining dynamic and temporal networks, and understanding the roles of individuals during the information-dissemination process.

This thrust builds strongly on our preliminary work. In particular, we have considered the problem of identifying controversial discussions on social media [19]. Our methods are designed to be applicable

to topics in any domain (e.g., political, economical, or cultural), and in a general setting. We have also addressed the question of devising a measure of controversy, which is applicable to any online discussion. More recently we devised a new study aimed to understand the phenomenon of echo chambers in social media, and to characterize distinct user types and their role in their local community [18].

In this project we will extend significantly our problem representations so as to obtain a deeper analysis and develop much more advanced knowledge-discovery methods. In particular, our existing work relies on analyzing plain graphs with no additional information regarding text or time. In REBOUND we will incorporate text information into the mining process. We will develop methods to extract semantic annotations from the textual data, such as (i) agreement between users; (ii) political leaning of users or content items; (iii) stance with respect to topics; and so on. We will use this information to enrich our graph models, provide novel formulations for mining semantically-annotated graphs, and then design efficient algorithms.

As an example, we will consider discussion graphs where edges are annotated with agreement ('+') or disagreement ('—'), and thus, represented as *signed graphs*. In this setting we will consider the problems of discovering alliances (dense subgraphs with many '+' and few '—' edges) or conflicting communities (sets of vertices with many '+' and few '—' within sets and many '—' and few '+' across sets). These are novel problems related to dense-subgraph discovery [32, 36, 37, 38] and correlation clustering [7, 21].

We will also leverage the temporal dimension of the data and study formulations that allow to model and discover bias, polarization, and conflict in the setting of temporal networks. This thread will make strong connections with our previous work on temporal networks [33, 34, 35] Finally we will consider problem representations related to motif discovery [5, 10] and role mining [3] in the context of REBOUND.

Research thrust 2: EXPLORATION

In this research thrust we will develop methods to help users understand better the global information landscape with respect to topics of their interest, visualize stories with annotations about different viewpoints, and increase their awareness about existing bias in the content they consume.

The underlying idea here is to allow users to visualize their content consumption so that they become aware of possible biases in their "news diet." At the same time we want to provide them the capability to visualize a more wide spectrum of information regarding a topic, and support them explore alternative viewpoints. Our preliminary work in this thrust includes a novel method for inferring political-leaning scores for user accounts and content sources in social media, in a completely unsupervised manner [23]. Those scores can be used in a number of different applications involving information-retrieval tasks or content visualization and exploration.

In REBOUND we will push this direction significantly beyond the state of the art. We will develop methods that produce comprehensive and understandable summaries of topics of interest discussed in online media. Our story summaries will contain information regarding the timeline of relevant events, important actors, entities, or concepts involved in the story, key persons participating in the discussion, and pointers to informative content. The summaries we will produce will allow users to understand (i) how polarized is a given topic; (ii) what is the spectrum of opinions about the topic; (iii) who are the most authoritative actors supporting a given opinion; (iv) what is best available content for that opinion; (v) what is the bias/position of media outlets regarding a topic; and so on. In addition, our framework will allow users to visualize their own position in the opinion space based on the content they have been exposed, their interaction with other users, and the content they have shared.

Finally, leveraging our framework for exploration and interaction we will develop methods to learn the probability that a certain user is interested in exploring a given information item. These probabilities will be used to optimize relevance of recommendations in the third thrust.

Research thrust 3: RECOMMENDATION

In this research thrust we will develop recommendation algorithms that aim to increase the overall diversity in the network and balance information exposure of conflicting views. We will also investigate the effect of different features to the willingness of the users to explore viewpoints that conflict their opinion.

Again, this thrust is strongly based on our preliminary results. In one of our most recent works we considered the problem of selecting a set of representative content items (such as, news articles with varying opinions on a given topic) to summarize a larger set of items [29]. This is a standard clustering problem, but in addition we require to select representative items that are close to each other, and thus, they are not polarized. For the resulting problem we are able to provide scalable algorithms with approximation guarantees and show practical relevance in the application scenario of selecting non-polarized representatives.

Additionally, we have studied the setting where we ask to make relevant recommendations to users, which however have positive externalities to the whole network. In particular, we have addressed the problems of balancing information exposure of different viewpoints in the network [17], and maximizing network diversity, where diversity is measured in different ways [4, 27].

In REBOUND we will advance the state of the art in many directions. First, we will enrich our networks with semantic annotations and will address recommendation problems in more realistic models. We will relax our assumptions and we will consider settings with multiple viewpoints, continuous polarity scores, and more complex information-propagation models. We will study novel recommendation problems, and will seek to resolve open problems in our current problem formulations.

Section c. Timeliness, feasibility, and contingency planning

As it is becoming apparent that social-media platforms have many adverse consequences, including bias, polarization, conflict, and creation of information silos, and as stakeholders are raising their voices to get those problems fixed, achieving viable solutions will require cooperation among companies, policy makers, and scientists in multiple disciplines, including academic computer scientists. In this respect, the project is extremely timely.

The PI is in a unique position to accomplish the goals of the project. His profile brings together theoretical work on algorithms design, with development of practical data-mining methods, and strong emphasis on applications. The practical aspect is further enhanced by a six-year experience in industrial research (Yahoo! Research) and a network of collaborators in big social-media companies. The PI is leading a research team of seven people, has been successful in international recruitment, and his doctoral students have acquired internship positions in companies like Google, Facebook, Amazon, LinkedIn, etc. Recent graduates are currently working in Google, EPFL, and University of Helsinki (assistant professor).

The PI and his research team have pioneered a line of work on the project theme and have obtained several preliminary results [4, 16, 17, 19, 23, 27]. The objective of this project is to advance the state of the art in this important topic on multiple fronts. In particular, we aim to work on more realistic models, achieve stronger algorithmic results, and resolve open problems. We also aim to enlarge and strengthen the team with researchers of complementary skills, and establish collaborations with industrial partners and scientists in other disciplines. Without the support of this project it will not be possible to accomplish these ambitious goals.

The research team has strong experience in collecting data from platforms like twitter and reddit, and much of our recent work has built on extensive data gathering. In addition, we will seek to establish collaborations with international online-media companies so that we can test our methods on real-world scenarios. We will also will seek collaboration with local media companies in Helsinki, such as Yle (Finland's national public broadcasting company) and Sanoma (a leading media group in the Nordic countries).

Risks. REBOUND sets ambitious goals and has high potential, but at the same time it contains significant risks. One of our risks is the possibility that a majority of people prefer to receive biased content and are not interested to explore conflicting viewpoints. This would be an interesting finding for the project. In this case, our knowledge-discovery methods will still be useful, but many of the the tools we will build they will be of interest to a smaller population. A second risk is that we will not be successful in devising algorithms with quality guarantees for the most challenging problems. In this case, we will make simplifying assumptions and focus on devising heuristic methods. A third risk is that we will not be able to establish high-profile industrial collaborations. In this case, we will perform our analysis on datasets we will collect, or public datasets, and will evaluate our algorithms using user studies.

References

[1] L. Adamic and N. Glance. The political blogosphere and the 2004 us election: divided they blog. In *Workshop on Link Discovery: Issues, Approaches and Applications (LinkKDD)*, pages 36–43, 2005.

[2] J. An, D. Quercia, M. Cha, K. Gummadi, and J. Crowcroft. Sharing political news: the balancing act of intimacy and socialization in selective exposure. *EPJ Data Science*, 3(1):12, 2014.

[3] A. Arockiasamy, **A. Gionis**, and N. Tatti. A combinatorial approach to role discovery. In *Proceedings of the IEEE 16th International Conference on Data Mining (ICDM)*, pages 787–792, 2016.

[4] C. Aslay, E. Galbrun, **A. Gionis**, and A. Matakos. Maximizing the diversity of exposure in a social network. 2018.

[5] C. Aslay, A. Nasir, G. De Francisci Morales, and **A. Gionis**. Mining frequent patterns in evolving graphs. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM)*, 2018.

[6] E. Bakshy, S. Messing, and L. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

[7] N. Bansal, A. Blum, and S. Chawla. Correlation clustering. *Machine learning*, 56(1-3):89–113, 2004.

[8] P. Barberá, J. Jost, J. Nagler, J. Tucker, and R. Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542, 2015.

[9] E. Bozdag. Bias in algorithmic filtering and personalization. *Ethics and information technology*, 15(3):209–227, 2013.

[10] M. Coletto, K. Garimella, **A. Gionis**, and C. Lucchese. A motif-based approach for identifying controversy. 2017.

[11] M. Conover, J. Ratkiewicz, M. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. Political polarization on twitter. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2011.

[12] G. Cormode, M. Garofalakis, P. J. Haas, C. Jermaine, et al. Synopses for massive data: Samples, histograms, wavelets, sketches. *Foundations and Trends® in Databases*, 4(1–3):1–294, 2011.

[13] A. Edwards. (how) do participants in online discussion forums create "echo chambers"?: The inclusion and exclusion of dissenting voices in an online forum about climate change. *Journal of Argumentation in Context*, 2(1):127–150, 2013.

[14] L. Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1962.

[15] D. Frey. Recent research on selective exposure to information. *Advances in experimental social psychology*, 19:41–80, 1986.

[16] K. Garimella, G. De Francisci Morales, **A. Gionis**, and M. Mathioudakis. Reducing controversy by connecting opposing views. In *ACM International Conference on Web Search and Data Mining (WSDM)*, pages 81–90, 2017.

[17] K. Garimella, **A. Gionis**, N. Parotsidis, and N. Tatti. Balancing information exposure in social networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 4663–4671, 2017.

[18] K. Garimella, G. D. F. Morales, **A. Gionis**, and M. Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the WebConference (WWW)*, 2018.

[19] K. Garimella, G. D. F. Morales, **A. Gionis**, and M. Mathioudakis. Quantifying controversy on social media. *ACM Transactions on Social Computing*, 1(1):3, 2018.

[20] E. Gilbert, T. Bergstrom, and K. Karahalios. Blogs are echo chambers: Blogs are echo chambers. In *International Conference on System Sciences*, pages 1–10, 2009.

[21] **A. Gionis**, H. Mannila, and P. Tsaparas. Clustering aggregation. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 2007.

[22] M. Grömping. "echo chambers" partisan facebook groups during the 2014 Thai election. *Asia Pacific Media Educator*, 24(1):39–59, 2014.

[23] P. Lahoti, K. Garimella, and **A. Gionis**. Joint non-negative matrix factorization for learning ideological leaning on twitter. In *Proceedings of the ACM International Conference on Web Search and Data Mining*

(WSDM), pages 351–359, 2018.

[24] Q. V. Liao and W.-T. Fu. Can you hear me now?: mitigating the echo chamber effect by source position indicators. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, pages 184–196, 2014.

[25] Q. V. Liao and W.-T. Fu. Expert voices in echo chambers: effects of source expertise indicators on exposure to diverse opinions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2745–2754, 2014.

[26] C. G. Lord, L. Ross, and M. R. Lepper. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, 37(11):2098, 1979.

[27] A. Matakos and **A. Gionis**. Tell me something my friends do not know: Diversity maximization in social networks. 2018.

[28] S. Munson, S. Lee, and P. Resnick. Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*, 2013.

[29] B. Ordozgoiti and **A. Gionis**. Consensus k-median. *Submitted for publication*, 2018.

[30] E. Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011.

[31] W. Quattrociocchi, A. Scala, and C. R. Sunstein. Echo chambers on facebook. 2016.

[32] P. Rozenshtein, A. Anagnostopoulos, **A. Gionis**, and N. Tatti. Event detection in activity networks. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1176–1185, 2014.

[33] P. Rozenshtein and **A. Gionis**. Temporal pagerank. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 674–689, 2016.

[34] P. Rozenshtein, **A. Gionis**, A. Prakash, and J. Vreeken. Reconstructing an epidemic over time. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1835–1844, 2016.

[35] P. Rozenshtein, N. Tatti, and **A. Gionis**. Finding dynamic dense subgraphs. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 11(3):27, 2017.

[36] M. Sozio and **A. Gionis**. The community-search problem and how to plan a successful cocktail party. In *Proceedings of the ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 939–948, 2010.

[37] N. Tatti and **A. Gionis**. Density-friendly graph decomposition. In *24th International Conference on World Wide Web (WWW)*, pages 1089–1099, 2015.

[38] C. Tsourakakis, F. Bonchi, **A. Gionis**, F. Gullo, and M. Tsiarli. Denser than the densest subgraph: Extracting optimal quasi-cliques with quality guarantees. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 104–112, 2013.

[39] K. Wallsten. Political blogs and the bloggers who blog them: Is the political blogosphere and echo chamber. In *American Political Science Association's Annual Meeting*, pages 1–4, 2005.