The Emergence and Collapse of Knowledge Boundaries

David A. Broniatowski [Member, IEEE] and **Christopher L. Magee**

Abstract

The dynamics of knowledge transfer is an important topic for engineering managers. In this paper, we study knowledge boundaries – barriers to knowledge transfer – in groups of experts, using topic modeling, a natural language processing technique, applied to transcript data from the U.S. Food and Drug Administration's Circulatory Systems Advisory Panel. As predicted by prior theory, we find that knowledge boundaries emerge as the group faces increasingly challenging problems. Beyond this theory, we find that knowledge boundaries cease to structure communications between communities of practice when the group's expert ability is insufficient to solve its task, such as in the presence of high novelty. We conjecture that the amount of expert knowledge that the group can collectively bring to bear is a determining factor in boundary formation. This implies that some of the factors underlying knowledge boundary formation may aid – rather than hinder – knowledge aggregation. We briefly explore this conjecture using qualitative exploration of several relevant meetings. Finally, we discuss implications of these results for organizations attempting to leverage their expertise given the state of their collective knowledge.

Index Terms

Group decision-making; Network analysis; Knowledge boundaries; Natural-language processing

I. Introduction

Engineering and technology organizations frequently employ interdisciplinary groups of experts to solve complex problems. These group members, representing multiple communities of practice (e.g., professional subspecialties or engineering fields in which members share common training and experiences), are expected to combine their knowledge to reach a solution that no group member could achieve individually [1–4]. However, knowledge transfer between group members from different communities of practice may be impeded in systematic ways. These impediments, known as "knowledge boundaries", [5] stem from differences in perception and interpretation of technical and other information [2, 6–10].

Previous research has largely focused on the mechanisms by which group members overcome knowledge boundaries, such as by using boundary objects and knowledge management systems [1, 5, 11–13], dialogue-based approaches [9, 14], or organizational solutions [15]. The boundaries themselves – and the state of the underlying expert ability – have apparently been presumed fixed.

This assumption seems especially problematic for the most innovative and interdisciplinary frontiers of knowledge. Therefore, we make a novel conjecture: that knowledge boundaries "collapse", or cease to structure patterns of communication, when groups of experts are faced with an extremely novel problem to solve.

The ability to adjudicate between theories about knowledge boundaries has been impeded by the fact that knowledge boundaries, themselves, have not been directly observed. In this paper, we use a new method to observe and compare knowledge boundaries. Specifically, we use a natural language processing (NLP) technique to quantify knowledge boundaries, enabling us to examine their effect on decision-making in domains of uncertain knowledge. This technique draws upon transcript text data to enable an analysis that is guided by the specific topics of the discourse (using topic models [16]). We are therefore able to incorporate contextual factors while still enabling an analysis that can generalize across these contexts. Furthermore, our method is extensible to the analysis of any group for which a text transcript exists, enabling results that may generalize beyond a small number of case studies. Thus, this research technique can be applied broadly – when texts exist – to problems where multiple sources of expertise are required to evaluate a technological artifact or situation. Examples include interdisciplinary R&D, multidisciplinary professional service firms, accident inquiries, and other instances of decision-making under uncertainty such as the FDA approvals studied empirically in this paper.

Beyond the application of these methodological contributions, we conduct qualitative explorations of several meetings to provide additional support for the "boundary collapse" conjecture. Our findings are especially important considering research showing that knowledge integration – the result of successfully overcoming knowledge boundaries – has been shown to improve performance on technical projects [17].

II. Literature Review

We examine two competing approaches to the problem of knowledge integration in groups of experts. These approaches make different predictions when groups are faced with situations that are so novel that their members' specialized knowledge does not apply. One approach argues that group members may use their expertise in inappropriate ways leading to a so-called "competency trap" [18]. (By expertise, we mean the use of intuitive but highly informed thinking in several fields, e.g., [19–21]). The other approach underlies our "boundary collapse" conjecture.

On one hand, Carlile [5] argues that expertise can create barriers to learning due to differences in interpretation that arise as novelty increases. These differences arise because group members draw upon different sources of knowledge associated with prior training and membership in different communities of practice. On the other hand, a considerable body of knowledge indicates that experts perceive situations differently than do novices (e.g., [22–26]). Experts can learn from one another to solve problems in novel situations [27–30], especially if group members have well-defined roles, responsibilities, and a strong sense of team orientation [20]. Furthermore, groups of experts in a novel situation sometimes arrive

at a common metaphor to aid understanding [9, 14, 31, 32]. These competing theories are described below.

A. Knowledge Boundaries

Groups of experts must learn from one another if they are to pool knowledge when solving complex problems. Despite a desire to do so, group members' prior knowledge can make this learning "laborious, time consuming, and difficult" [33, 34]. Often, the knowledge to be communicated is "tacit" [19, 35], meaning that it cannot be explicitly communicated through speech. Rather, it must be shown to, and practiced by, the recipient. For example, a surgeon may be unable to explain a complex surgical procedure to a biostatistician to construct a predictive model of success rates because the specifics of the procedure are encoded in muscle memory. The difficulties communicating this "deep knowledge" constitute a major barrier to learning [2, 7, 14, 36–38].

Such difficulties depend both on the nature of the knowledge to be communicated and the prior expertise of the communicators. Carlile defined three types of "knowledge boundaries" that arise as novelty increases, namely syntactic, semantic, and pragmatic boundaries. Syntactic boundaries occur in the least complex and novel scenarios, followed by semantic boundaries, and then pragmatic boundaries [1].

1) Syntactic Boundaries—Syntactic boundaries arise when two individuals do not share a common set of terms, such as a natural language, to describe a problem [39]. For example, American English speakers may refer to a device as an "elevator" whereas British English speakers may refer to the same device as a "lift". These boundaries are frequently overcome by establishing a common vocabulary when discussing a problem. Since the committees studied in this paper consist of members who share a common national culture, observation of syntactic boundaries is outside the scope of our analysis; nevertheless, we mention them for completeness

2) Semantic Boundaries—Even if group members share a common vocabulary, they may encounter semantic boundaries. These are impediments to learning that arise from differing interpretations of the same data or situation. These different interpretations are associated with assumptions, standards of rigor, or aims that differ between communities of practice whose members share common experiences and training [2, 3, 5, 7, 8, 10, 19, 40–43]. For example, different group member may have different ideas of what constitutes a "safe" or "effective" treatment based on their own prior experiences. Although grounded in tacit knowledge, semantic differences can be overcome by discussion that refers to the meanings of the terms used (e.g., "by safe, I mean that the patient's risk of death is less than 2%") [14, 44].

3) Pragmatic Boundaries—Often, new knowledge must be created in response to a novel situation [5]. Carlile theorizes that, instead of creating new knowledge, experts prefer to rely upon their existing knowledge leading to a "competency trap" [18]. Since different group members have different sources of prior knowledge, this can lead to differences in goals between communities of practice as each favors interpretations that are most consistent

with their prior experiences to the exclusion of others. For example, even though two group members might have the same definition of "safety", one group member might interpret a dataset in a medical setting as indicating that a certain treatment is unsafe whereas another might interpret the same data as indicative of a badly designed measurement that should be discounted in lieu of other evidence. Thus, Carlile [5] argues that experts may be unable to learn from one another because of deep assumptions rooted in their prior experience.

B. Boundary Collapse

Scholars of expertise and intuition are more optimistic regarding the capacity for groups of experts to make decisions in unfamiliar conditions. Experts are able to operate successfully in novel situations, and recognize when a situation is familiar or atypical. In atypical situations, experts spend most of their time assessing their own ability to contribute [46–49]. Nevertheless, there are some very novel situations where experts' abilities may no longer be sufficient [50]. Experts can recognize such situations as atypical and will not rely on expert knowledge. For example, Chase & Simon [22] found that chess masters were indistinguishable from novices when confronted with random (i.e., novel) chessboard configurations.

Prior work has found that groups of experts tend to base their decisions on shared common (non-expert) knowledge when facing the most novel situations. For example, Faraj and Xiao [51] found that in the most difficult cases, members of trauma teams from different communities of practice did share knowledge, but only general knowledge. Furthermore, Lamont [52] has observed that experts from the same community of practice may make decisions that diverge significantly, despite common expertise, in a manner that is not consistent with Carlile's definition of pragmatic knowledge boundaries. Finally, Majchrzak et al. [9] found that group members with weak social ties who faced a novel task under time pressure did jointly create a solution, but this solution did not make use of inapplicable expert knowledge. In each of these cases, experts seemed to be able to diagnose their own specialized knowledge as irrelevant to the problem at hand, and decisions were instead based upon common knowledge. This bears some similarity to the "hidden profile effect" (e.g., [53–55]) – a classical finding in social psychology pertaining to groups of novices – in which common knowledge strongly drives decisions whereas even relevant unique knowledge remains unshared because it is not socially validated.

C. Representing Knowledge Boundaries

We aim to adjudicate between different predictions regarding the impact of extreme novelty on knowledge boundary formation. On one hand, Carlile's framework predicts that pragmatic knowledge boundaries will lead to a "competency trap". On the other hand, the expertise literature suggests that boundaries will "collapse", leading to the sharing of common knowledge, but not expert knowledge. To learn which concept is most applicable, we must observe the emergence and collapse of knowledge boundaries.

Our approach emphasizes convergent-divergent validity. By testing the performance of our technique where both approaches agree, we establish the construct validity of our approach. Once established, we use this technique to adjudicate between these two approaches.

1) Representing Communities of Practice—Knowledge transfer occur within "cohesive" networks in which members share third-party ties or similar mutual connections [56–59]. Communities of practice provide this cohesion via common standards of evaluation and rigor. These communities are frequently described as overlapping networks of people who use the same tools to accomplish similar tasks across several different contexts [60–62]. For example, trained electrical engineers can apply their expertise to domains as different as biomedical device design and aerospace system design.

Brown and Duguid [61] indicate that social network analysis is an appropriate formalism with which to describe communities of practice. Social networks are made up of nodes and edges, where a node represents an entity, and an edge is a flow of some quantity between nodes [63]. When representing communities of practice, a node represents a group member and an edge represents the transfer of knowledge between members [59]. Since common standards of evaluation and rigor are associated with common language and jargon, we measure the extent to which group members transfer knowledge by determining whether they share a common topic of conversation [64]. Specifically, we use the technique described in [65] to trace knowledge flow by automatically constructing networks from meeting transcript texts. The absence of an edge between conversing group members indicates that they do not share a common topic of conversation.

2) Representing Semantic Boundaries—Each edge in a network represents a shared topic of conversation. Communities of practice are therefore represented by groups that are internally cohesive -- i.e., densely linked within a community. Furthermore, if these groups are sparsely linked, or even disconnected, across communities (see Fig. 1), this implies the existence of semantic knowledge boundaries.

3) Representing Pragmatic Boundaries—Pragmatic knowledge boundaries are defined by differences in practice (i.e., behavior) given access to the same information. Pragmatic knowledge is tacit, so an expert may only be able to communicate a preference for a particular decision (e.g., a system design option, or voting outcome) rather than an explicit rationale. In a network exhibiting pragmatic boundaries, we expect group members who do not vote the same way to be disconnected. Pragmatic knowledge boundaries therefore imply shared voting behavior, with similar behavior among group members that are connected, but different behavior otherwise (see Fig. 2).

D. Research Hypotheses

These representations enable us to test hypotheses in a quantitative, replicable manner. In our networks, each node represents a member of a decision-making group. Two nodes are linked if they share a topic. Conversely, a knowledge boundary that impedes sharing is represented by the absence of a link between two nodes.

By definition, semantic boundaries are present where individuals do not share at least one common topic of conversation [64]. Furthermore, semantic boundaries should exist between communities of practice. Therefore, in a network where semantic boundaries exist, we expect individuals within the same community-of-practice to be connected, with fewer connections between such communities. Thus our first hypothesis:

Hypothesis 1: Communities of practice will be cohesive – i.e., links (indicating shared topics of conversation) will be more likely to form between members of the same community than between members of different communities.

Similarly, pragmatic boundaries exist imply that individuals who do not share the same goals will not communicate with one another. Thus, we expect individuals who do not vote the same way to be disconnected. Thus our second hypothesis:

Hypothesis 2: Voting groups will be cohesive – i.e., links (indicating shared topics of conversation) will be more likely to occur between group members who vote the same way than between members who vote differently.

Furthermore, Carlile's hierarchical framework implies that pragmatic boundaries should only occur when semantic and syntactic boundaries are already present. Meetings with cohesive voting groups should therefore have cohesive communities of practice. This leads us to formulate two versions of our third hypothesis:

Hypothesis 3a: Cohesive voting implies cohesive communities of practice.

Hypothesis 3a means that one cannot have a pragmatic boundary without also having a semantic boundary.

Hypothesis 3b: Cohesive communities of practice do not imply cohesive voting.

Finally, semantic boundaries can occur even when pragmatic boundaries do not. Consequently, meetings without pragmatic boundaries (including meetings where the group reached consensus or had only one voter in the voting minority) should still display semantic boundaries.

Each of these hypotheses, although derived from Carlile's framework, are nevertheless also consistent with the expertise-driven approach. In the following section, we describe a situation where the predictions of these two approaches differ.

E. The Boundary Collapse Conjecture

Carlile's theory predicts that pragmatic boundaries prevent learning between group members in the most novel situations because group members reuse inapplicable knowledge that varies between communities of practice. If this is true, we should see cohesive voting groups in these situations. In contrast, we conjecture that group members recognize the limits of their expertise, leading knowledge boundaries to "collapse" in the most novel situations. This means that communities of practice and voting groups will not be cohesive in the presence of novelty, because inapplicable knowledge is not deployed; rather, decisions are made upon the basis of shared general, rather than specialized, knowledge.

Conjecture: In highly novel situations, neither specialty groups nor voting groups will be cohesive.

III. Methodology

In order to measure the emergence and collapse of knowledge boundaries, we used natural language processing techniques to analyze transcripts of expert group meetings. We studied

37 FDA advisory panel meetings to determine whether common patterns of knowledge transfer were found across groups.

A. Case Selection

We analyzed meeting transcripts from the U.S. Food and Drug Administration's (FDA) Circulatory Systems Devices Panel. We chose this panel because it had a large number of meetings recorded. Our sample is the full set of meetings held between 1997 and 2005 where panel members voted regarding the Pre-Market Approval (PMA) of a medical device. Although our analysis was restricted to this panel, there is no reason why it could not be extended to other panels, agencies, or communities of practice in general. Devices brought to these committees for review are generally those which the FDA does not have the "in-house" expertise to evaluate. As such, the devices under evaluation by the committees are likely to be the most innovative, and those facing the most uncertainty or novelty. This is similar to the conditions faced by many cross-functional expert groups, which are often convened to make recommendations on questions with a similar degree of uncertainty.

We extracted the names, communities of practice – i.e., medical specialties (e.g., surgeons, cardiologists, radiologists, electro-physiologists, statisticians, etc.), and votes of the panel members for each of the 37 meetings studied. The verbatim text of each meeting transcript was then divided into "utterances," where each utterance was a fixed paragraph of text spoken by an identified panel member as denoted by the meeting's court recorder. Utterances are sequential, and therefore are used to denote the order in which panel members speak. Utterances can range in length from one or two words (e.g. "thank you") to several sentences (about 100 words), but most are between 50 and 75 words in length.

1) Why study meeting transcripts?—We rely upon text data because dialogue is one of the primary mechanisms by which knowledge is created and transferred [14]. In addition, functioning within a community of practice requires the use of specialized language and terminology [1, 46, 49]. Dialogue text is therefore ideal data with which to study knowledge transfer within expert groups. Since dialogue is neither regular nor fully deterministic in its structure, probabilistic analysis is required.

Prior work [16, 64–67] has shown how Bayesian topic models could be used to infer probabilistic topics, containing semantic-level representations of a text. Two group members who share knowledge are likely to discuss the same topic (although they may not necessarily agree). We therefore used probabilistic topic models to infer semantic content from meeting transcripts.

B. Bayesian Topic Models

Approaches based on Bayesian inference provide an ideal platform to study semantic relationships [64]. Of particular interest are topic-modeling approaches to studying social phenomena in various contexts. A topic model seeks to divide a corpus into a small number of discrete, but semantically-coherent, topics (see Table I). Quinn et al. [68] argue that social scientists should use topic models because they impose a minimal number of transparent

assumptions due to their mathematical underpinnings. A full mathematical description of our Bayesian topic modeling technique is given in the Appendix.

We extracted semantically-coherent topics from meeting transcripts in order to examine semantic cohesion. The approach presented here may be viewed as an extension of "latent coding" – one type of formal content analysis prevalent in the social sciences. The most important limitations of latent coding, and other hand-coding methods, are the inability to scale to large numbers of documents. This limitation stems from a dependence on the coder's knowledge, leading to interrater reliability concerns. Furthermore, hand-coding is labor-intensive, often requiring teams of trained coders. One example of a latent coding scheme might aim to map specific words or phrases to pre-defined categories. Coders would then count the number of word in each category that a particular speaker uses. Variance between coders would arise from words or phrases that do not clearly fit in a pre-defined category. This is a likely occurrence in analyses of expert teams due to the context-specific and highly technical nature of the meetings analyzed: identifying words that might be important is difficult *a priori*. Additionally, such coding schemes are subject to confirmation bias by researchers seeking specific theoretical constructs. The motivation behind using a computational approach is therefore to create a method that is automatic, repeatable and consistent – and therefore replicable by other researchers.

C. Network Construction

Having grouped words into topics, we next generated a network representing the flow of knowledge between speakers for each meeting. Because of the probabilistic nature of the Bayesian topic models, 200 samples were taken for each transcript to enable averaging. Given one such sample, a speaker, X_i, has a probability distribution over each topic, $P(Z|X_i)$, for every one of the 200 samples generated above. A pair of speakers is connected by an edge within a given sample if their joint probability of discussing the same topic is greater than chance

$$P(X_i \cap X_j) = \sum_t^T P(Z = z_t | X_i) * P(Z = z_t | X_j), \quad (1)$$

where T is the total number of topics. Examining all 200 samples, [69] determined that a given pair of speakers are considered linked in a given aggregate knowledge network if they are connected in at least 125 of the 200 samples. We generated one such aggregate knowledge network for each meeting transcript.

1) Measuring Cohesion—Our hypotheses refer to cohesive communities of practice (e.g., medical specialties) and cohesive voting groups. We defined metrics for these cohesion values as follows: A network has cohesive specialties when its links are mostly between members of the same specialty.

We defined *raw specialty cohesion* (RSC) as the number of links between members of the same specialty, l_s, divided by the total number of links, l_t, in the network:

$$\text{RSC} = \frac{l_s}{l_t}. \quad (2)$$

In order to enable comparison across several meetings, the raw specialty cohesion values of 1000 random graphs were then compared to the raw specialty cohesion value of the knowledge network. 1000 random graphs were used because this number was empirically determined to generate stable results. We defined s*pecialty cohesion percentile* (SCP) as the total number of random graphs with a RSC value that is less than the RSC value of the network derived from the AT model algorithm.

$$f(i) = \begin{cases} 0, & \text{RSC(empirical)} < \text{RSC}(i) \\ 1, & \text{RSC(empirical)} \geq \text{RSC}(i) \end{cases} \quad (3)$$

This number was then divided by 1000 so that it lies between 0 and 1 (i.e., it is converted to a percentile score).

$$\text{SCP} = \frac{\sum_{i=1}^{1000} f(i)}{1000} \quad (4)$$

For example, the idealized graph depicted in Fig. 1 has a specialty cohesion value of 1.0. The graph in Fig. 3 is an example of one of the FDA meetings analyzed for which members of the same specialty tend to be linked to one another, whereas members from different specialties tend not to be linked. It has a SCP value of 0.97.

Meetings with a SCP that is greater than 0.5 show more intra-specialty links than would be expected by chance, whereas meetings with a SCP that is less than 0.5 display more links between specialties than would be expected.

We defined *raw vote cohesion* and *vote cohesion percentile* the same way: raw vote cohesion (RVC) is the number of links between members who voted the same way, divided by the total number of links, l_t, in the network, whereas vote cohesion percentile (VCP) is the total number of random graphs with a RVC value that is less than the RVC value of the network derived from the AT model algorithm. VCP was calculated for the 11 meetings in which there was a voting minority of at least two members (for the remaining 26 meetings, vote cohesion is undefined since no links can exist within a voting minority of size zero or one). VCP also lies between 0 and 1. For example, the graph in Fig. 2 has VCP of 1.0 and the graph in Fig. 4 has VCP of 0.99.

IV. Results

We calculated SCP values for each of the 37 meetings in our sample. A 1-sided Kolmogorov-Smirnov test, *p=0.0045, n=37*, shows that SCP is significantly higher than

expected by chance among these 37 meetings. Consequently, we reject the null hypothesis in favor of Hypothesis 1.

We next calculated VCP values for the subset of 11 meetings in our sample in which there were at least two voting minority members. A 1-sided Kolmogorov-Smirnov test, *p=0.02, n=11*, shows that VCP is significantly higher than expected by chance among these 11 meetings. Consequently, we reject the null hypothesis in favor of Hypothesis 2.

Finally, we examined the relationship between VCP and SCP for the 11 meetings in which VCP was defined (see Fig. 5).

These two quantities are strongly associated, *Spearman rho=0.79, p=0.006*. In our sample, there are no points in the upper-left or lower-right portions of the graph: the absence of points in the upper-left indicates that there are no meetings with high VCP (indicative of pragmatic boundaries) that do not also have high SCP (indicative of semantic boundaries). Consequently, we accept hypothesis 3a that pragmatic boundaries entail semantic boundaries.

For the remaining 26 meetings in which VCP was not defined, SCP is still significantly higher than we would expect due to chance *p=0.0093, n=26*, suggesting that semantic boundaries do occur in the absence of pragmatic boundaries. Thus, we accept hypothesis 3b.

A. Examining the Boundary Collapse Conjecture

These hypotheses demonstrate the ability of our method to analyze knowledge boundary formation in groups of experts. They also support accepted theory in this area. However, our hypotheses do not explain the absence of data points in the lower-right quadrant of Fig. 5. Our data show that low VCP is associated with low SCP. In the seven non-boxed meetings in Fig. 5, there is a sizable voting minority, yet members of these different voting groups are not likely to be linked; nor are members of the same medical specialty linked. These results are consistent with the boundary collapse conjecture. We also conjectured that the results might be associated with high novelty; nevertheless, these data provide no explicit information about the meeting's relative novelty or the relevance of panel members' knowledge. We therefore explored several meetings in more depth so that we could gain insight into the relationship between specialty and vote cohesion percentiles under varying conditions of novelty. Meetings were rated as "low", "moderate", or "high" novelty by one of the authors (DAB). Although these descriptions are not meant to be a formal qualitative analysis, we feel that they provide important context for the quantitative analysis presented above. We examined meetings with varied vote cohesion and specialty cohesion findings.

1) The Medtronic Model 7250 Jewel® AF Implantable Cardioverter Defibrillator System (12/5/2000)—The graph for this device is found in Fig. 3. This device had previously been approved for sale by the FDA: the purpose of this meeting was to broaden its use to a wider array of illnesses [69]. This was well within the range of the panel members' expertise, and sufficient knowledge was apparently easily brought to bear despite semantic boundaries. The panel unanimously voted to approve the device subject to conditions, including language that limited the device's use to a slightly more specific range

of illnesses. This change in language represents a concrete example of knowledge transformation used to overcome pragmatic boundaries between various panel members – a panel member who had considered not approving the device did so subject to this condition. This low-novelty meeting had a high specialty cohesion percentile value of 0.968, indicating strong semantic boundaries and the deployment of expertise. Thus, a potential pragmatic boundary was overcome to build consensus.

2) The Eclipse Holmium TMR Laser (10/27/1998)—The Eclipse device was the first Holmium (solid state) TMR laser to be approved by the FDA. Furthermore, the safety and effectiveness of the associated surgical procedure had only recently been demonstrated. Finally, there were concerns regarding the quality of the data resulting from the clinical trial. There was one panel member who voted against device approval, citing these data concerns, whereas a member of the voting majority summarized his reason for approving the device as follows: "I think this is consistent with what has been done with the other similar device and I think is it probably appropriate for the FDA to be consistent," [69] indicating low novelty because the panel's previous experience with similar devices could be brought to bear. The SCP value of this meeting was high – 0.975 – indicating strong semantic boundaries and the deployment of expertise. Consequently, this case provides an example of a meeting in which high SCP is associated with a lower level of novelty than would have been available had the procedure itself not been previously examined by the FDA panel.

3) The Eclipse Holmium PMR Laser (7/9/2001)—This device required a new clinical procedure whose mechanism was unknown resulting in questions regarding device efficacy. Ultimately, every cardiologist, and the panel's statistician, voted not to approve the device, whereas the two non-cardiologists – a surgeon and a specialist on the placebo effect – both voted in favor of device approval. Cardiologists' reasons for not approving the device reflected the role of novelty in driving panel members' decisions. The SCP value of this meeting was high – 0.959 – indicating strong semantic boundaries and the deployment of expertise. VCP was at its maximum value of 1.00, indicating strong pragmatic boundaries. This case thus provides an example of a meeting in which high SCP and high VCP are associated with moderately high novelty and several practice-driven perspectives on the same device's data – the pragmatic boundary and semantic boundaries overlap.

4) The Cryocath Freezor Cryoablation Catheter (5/6/2003)—The graph for this device is found in Fig. 4. The device's clinical trials failed to reach any of its specified data endpoints, precluding statistically valid conclusions regarding safety and efficacy. Electrophysiologists and surgeons on the panel drew upon post hoc analysis and expertise to infer how the device might perform in clinical practice. In a 6-3 vote, the panel recommended device approval. All electrophysiologists and surgeons voted in favor of the motion to approve the device; cardiologists and the statistician voted against the motion. This voting split reflected a dichotomy between clinical and statistical sources of expertise on the panel, and serves as an example of voting in line with pragmatic knowledge. The SCP value of this meeting was 0.996 indicating strong semantic boundaries and the deployment of expertise. The VCP value of this meeting was 0.987, indicating strong pragmatic boundaries. Thus, this case provides an example of a meeting in which high SCP and high

VCP are associated with a disagreement regarding which source of knowledge is more relevant – a pragmatic boundary associated with a semantic boundary.

5) The Abiocor Fully Implantable Replacement Artificial Heart (7/23/2005)—The device was novel – the first self-contained artificial heart system – promising the potential to revolutionize the practice of medicine. A device failure would almost certainly mean death for the patient. Furthermore, device implantation was a difficult surgical procedure with a long recovery time that would only be performed on those patients who would otherwise have died due to their illness. Due to the high-stakes nature of this device's approval, the FDA opted for a Humanitarian Device Exemption allowing a smaller amount of data to be gathered before its release to the market. The clinical trial for the device therefore only included 17 patients, and data was inconclusive regarding the device's safety and efficacy. Panel members' interpretations of clinical trial results differed, but these differences were not associated with specialized knowledge. Ultimately, in a 6–7 vote, the panel voted down a motion to approve the device. The SCP of this meeting was very low: 0.087 (suggesting that group members were explicitly searching for information from outside their professional specialties) indicating the collapse of semantic boundaries and that expertise was not deployed. The VCP of this meeting was 0.586 – near randomness which is associated with a VCP of 0.5 – indicating the collapse of pragmatic boundaries. In this situation of extreme novelty, when there is insufficient knowledge, existing communities-of-practice have little impact on communication or on voting.

6) The ACORN CorCap Cardiac Support Device (6/22/2005)—This device was extremely novel. It represented the first attempt to control heart enlargement by surrounding the organ with a mesh support sock during open heart surgery. The clinical trial associated with this device encountered several problems including large amounts of missing data, and problems with subjects who had become unblinded to their treatment. This led to difficulty interpreting device efficacy except in one endpoint – death – for which the device did not show a significant improvement when compared to the control group. Due to the novel nature of the device, clinical expertise could not be brought to bear. Ultimately, in a 9-4 vote, the panel accepted a motion not to approve the device. The SCP of this meeting was 0.494 – near randomness – indicating the collapse of semantic boundaries and that expertise was not deployed. VCP of this meeting was 0.277 (perhaps suggesting that most conversation involved disagreement), indicating pragmatic boundary collapse. As in Meeting 5, this meeting indicates that in situations where there is insufficient knowledge present, communication within communities-of-practice during the meeting is weak and has little impact on the vote, apparently because specialized relevant knowledge cannot be identified.

7) The Cordis CYPHER Sirolimus Drug Eluting Stent (10/22/2002)—This device was the first implementation of a very novel technology – drug-eluting stents (DES). The arrival of CYPHER was highly anticipated; one cardiologist panel member commented that "[t]his is going to revolutionize our profession" [69]. Much of the discussion during the panel meeting surrounded conditions of approval rather than whether or not the device should be approved, and the panel unanimously recommended device approval. DES became popular among physicians, quickly diffused throughout practice, and were often

implanted for off-label use. Four years later, in 2006, several studies presented results showing that patients with DES had a higher risk of adverse events than patients who had received standard (bare metal) stents [70–72]. The SCP for this meeting was 0.362 indicating the collapse of semantic boundaries, and the non-deployment of expertise, despite consensus. Since the device was approved unanimously, VCP was not defined. This case provides an example of a meeting in which low SCP is associated with a comparatively high level of novelty and a low degree of previous experience. Although all panel members agreed that the device should be approved, the low SCP might indicate that no expertise related to medical specialty was brought to bear on this decision. This indicates that in situations of extreme novelty; when there is insufficient knowledge, communication during the meeting is not strongly affected by existing communities-of-practice, and has little impact on the vote. In this particular case, widespread pre-existing enthusiasm about DES may have played a decisive factor in the device's approval.

Table II summarizes this discussion and its implications for the boundary collapse conjecture. Contrary to expectations, in situations where available knowledge is apparently sufficient to solve the problem (i.e., novelty is low), SCP is high. Semantic boundaries exist because specialized expertise from multiple perspectives is being brought to bear (recall that the FDA Advisory Panels' charters call for some degree of novelty) as in meetings 1 and 2, but major disagreement does not arise. As novelty increases, voting differences and pragmatic boundaries may emerge. These are coupled with semantic boundaries as in meetings 3 and 4, where VCP and SCP values are both high. Finally, where the existing expert ability is insufficient or inapplicable (i.e., when novelty is high), SCP decreases below 0.5 – below the level of randomness – as in meetings 5, 6, and 7. This is consistent with the behavior of a panel that is sharing general information across specialty boundaries – i.e., expert knowledge is not mobilized because it is not applicable. Where it is defined, VCP also decreases, indicating that voting behavior is unrelated to the topics of discussion. These results provide preliminary support for the boundary collapse conjecture.

Alternative explanations for the broad subject of what happens under conditions of high novelty include "competency traps" [18] and political processes resulting from knowledge that is "at stake" due to investments in obtaining expertise [1]. The transcripts offer no support for these alternative hypotheses, although they cannot be fully eliminated since such evidence might not be easy to discern.

V. Discussion

A. Methodological Implications

By inferring topics from meeting transcripts, we are able to observe patterns that are consistent with empirical and theoretical findings in the literature on organizations (e.g., [5, 6, 11, 12, 13–15]. Our technique therefore partially confirms existing theories of how knowledge boundaries arise [1, 9]. These results simultaneously indicate the validity of our method for studying knowledge boundaries in groups of experts. Finally, our technique is broadly applicable to any text transcript data source.

B. Theoretical Implications

Our results support the notion of boundary collapse. Consistent with Carlile's theory, we detected communication patterns consistent with the existence of semantic boundaries for most of the meetings in our sample. Additionally, we detected communication patterns consistent with the existence of pragmatic boundaries for those meetings in which the panel had voting minorities with at least two members. Furthermore, we found evidence that pragmatic boundaries entail pragmatic boundaries, but not vice versa. All of these findings are consistent with Carlile's framework.

Beyond Carlile's framework, we found evidence supporting our boundary collapse conjecture. In particular, the existence of a class of meetings that had low SCP and VCP values indicates that neither type of knowledge boundary is structuring interaction between group members. Our qualitative analysis provides additional insight into the underlying mechanisms. It appears that knowledge boundaries are not present when experts are able to recognize that their specialized knowledge does not apply. Under these circumstances, members' expertise plays a key role in helping them to recognize the limits of their knowledge. Values of SCP that are less than 0.5 are consistent with more communication between communities of practice than within communities of practice. In other words, group members may be engaging in search behaviors – looking for information from external sources. Similarly, low values of VCP indicate that group members did not vote in blocs reflecting their communities of practice.

This study is the first, to our knowledge, to empirically explore how knowledge boundaries emerge and collapse in a large number of groups of expert decision-makers outside of a laboratory setting. It is also the first to use a computational method to explore the emergence and collapse of knowledge boundaries wherever transcripts of decision-making groups are available.

Previous literature has presumed that knowledge boundaries remain fixed after they have arisen (e.g., [1, 15]), requiring that they be either traversed or transcended [9]. Our conjecture extends these theories with our finding that, when novelty is high, expertise effects (i.e., specialty- and vote-cohesion) decrease. Our results suggest that semantic boundaries are indicative of the deployment of specialized knowledge within a community of practice. Furthermore, pragmatic boundaries emerge at the cusp of expertise, where there is enough knowledge such that communities of practice can reach internal consensus, but there is little agreement between communities regarding how best to combine or deploy available knowledge. These findings are therefore especially relevant for the most innovative and interdisciplinary frontiers of knowledge.

C. Implications for Engineering Management

Engineering managers have long recognized the existence of a large class of problems, known as "wicked" [73], "messy" [74], or "swamps" [32], that resist solution because of conflicting goals, values, and sources of knowledge and expertise. Similarly, decision-makers may be influenced by "group think" [75], dominated by a small number of strong personalities, or subject to other social dynamics that can lead to an ill-informed outcome.

Sometimes, a group may not have the knowledge required to make a well-informed decision and should instead devote more time to researching the particulars of the problem. In other situations, stakeholders with valuable expertise to contribute may not be present because the need for their expertise may not have been recognized.

Franco [76] recently argued that "Soft OR" methods (or Problem Structuring Methods; PSMs, e.g., [77]). might be used to address these problems because they can generate "boundary objects" [13] –artifacts that help decision-makers to share information across knowledge boundaries. According to Franco's framework, these methods and the models they generate, are successful to the extent that they aid decision-makers to recognize, and ultimately overcome, these boundaries.

Our technique may be used to evaluate such interventions. Specifically, it can help analysts to determine whether knowledge boundaries are present and, if so, what kind.

By enabling the measurement of knowledge boundaries, our technique could feasibly be used to determine the efficacy of these techniques in helping a decision-making group effectively traverse these boundaries.

Our study contributes to a growing knowledge base (e.g., [76]) that suggests how engineering managers might use models and data to overcome knowledge boundaries. In situations with relatively low novelty, pragmatic boundaries might be overcome if the boundary object focuses group members on the specific areas of their disagreement. Under these circumstances, solutions are possible if group members share common values (e.g., prioritizing a device's safety over its efficacy). Practitioners can then guide the group towards a solution designed to address this specific area, e.g., when panel members limited device indications to specific, less-problematic, sub-populations in Meeting 1, and when stricter follow-up protocols were imposed in Meeting 2. For moderately novel situations, group members may draw analogies with their prior experience, such as when panel members compared PMR to TMR in Meeting 3, and cryoablation to heat ablation in Meeting 4, leading them to focus on the elements of the statistical data and model output that they believe are most representative. When different group members draw different conclusions because their expertise leads them to focus on different results within the dataset, pragmatic boundaries are present, and will usually prevent the emergence of consensus. Techniques designed to overcome these knowledge boundaries that are able to lead decision makers to agreement regarding values may be able to overcome these pragmatic boundaries.

Our technique has an especially important role to play in the most novel situations. If group members can recognize that their expertise does not apply in such cases, they are more likely to avoid a "competency trap," and inappropriately apply inapplicable expert knowledge. A technique that makes the uncertainty in these situations explicit may help experts to recognize the limits of their abilities, such as in Meetings 5 and 6 when panel members recognized that data and methodological limitations in the clinical trials precluded drawing clear conclusions. In these scenarios, our technique could help engineering managers

recognize the presence of deep uncertainty, motivating the need to seek new sources of expertise, or redefine the problem to something more tractable.

D. Limitations and Future Directions

Our study is limited to a sample of 37 FDA Circulatory Systems Advisory Panel Meetings, and therefore may not generalize to all expert groups. Nevertheless, our findings are based upon theory developed from across a wide range of domains. Thus our analysis serves to extend this theory into a novel context. In our study, we relied on meeting transcripts in order to examine the emergence and collapse of knowledge boundaries in expert groups. Our data do not account for several rich sources of information, e.g., body language, which might be available in a traditional field study. Furthermore, the nature of our data set is inherently observational and so we cannot control for panel membership, pre-existing relationships among panel members, etc. In addition, our analysis focused on one panel in one agency. Thus, an immediate extension would explore the extent to which our technique and results would apply in other group settings across government, industry, and academia including studies of real-world, but less expertise-based, groups. Finally, our technique is limited to those situations where transcript data are available. However, similar public transcripts exist for other decision-making bodies such as the Federal Open Market Committee (FOMC), political committees, deliberations of the board of directors for some publicly-traded companies, etc. Furthermore, several such transcripts may be found across a variety of contexts due, in part, to reporting requirements imposed by the Federal Advisory Committee Act (FACA) of 1972, which guarantees that transcripts of several committees may frequently be obtained as a matter of public record. In addition, transcript data of the sort required by our method has been collected for design teams in a straightforward manner by scholars of engineering design (e.g., [78]). Finally, the costs of transcribing such meetings continue to fall with the advent of crowdsourcing and other techniques that allow easy recruitment of online workers. Thus, we believe our technique to be applicable across a wide range of contexts. Thus, situations where there is no possibility to transcribe the dialogue between the experts, such as first responder teams, construction crews, etc., are unlikely to fall within the scope of what our technique can address – nevertheless, the underlying theory applies generally to cross-disciplinary teams. For example, Majchrzak et al. [9] built on Carlile's framework when examining cross-functional teams in high-stakes surgical settings. No technique is universally applicable, and data limitations are always present in empirical studies; however, the strength of our approach is that it extends the empirical possibilities for quantitative analysis of cross-functional teams.

Our technique might be useful in identifying and exploring the qualities of individuals who seem to be most adept at translating and transforming knowledge across semantic and pragmatic boundaries. One exciting area of future work would address how knowledge boundaries might serve a broader role than simply impeding the flow of knowledge between experts. For example, one might conjecture that semantic boundaries serve to provide a structure through which expertise can be aggregated. As such, they might aid, rather than hinder, group decision-making. Our results suggest that knowledge must be sufficient both to appropriately catalyze the formation of expert sub-groups while preventing the emergence of pragmatic boundaries that might cause a group to fracture.

VI. Conclusions

Our analysis of knowledge boundaries suggests that their behavior is not as simple as has previously been conceived. In particular, knowledge boundaries may cease to structure information flow in very novel situations. Group members therefore cease to rely upon their expert knowledge and, in their search for a solution, rely exchange only general information. In this way, they become indistinguishable from novices. The tools and techniques presented in this article provide a means to study these knowledge boundaries, and their behavior, directly, enabling a deeper understanding of the dynamics of knowledge transfer in decision-making groups than was possible before.

Acknowledgments

Appendix

Topic Models

Bayesian topic models [67] assume that each word is assigned to a discrete topic with a given probability. Each topic is assumed to be *exchangeable*, i.e., conditionally independent of each other topic [80]. This assumption allows the same word to occur in two different topics. A word is therefore modeled as having been drawn from a discrete probability distribution over topics.

Latent Dirichlet Allocation (LDA; [66]) is the most widely-applied, Bayesian topic model. In LDA, each *word* (w) is assigned to a *topic* (z) with a given probability that is inferred by the topic modeling algorithm. Each topic is therefore defined by a probabilistic distribution (ϕ) over all the words in the corpus (i.e., in one meeting). This distribution is *multinomial* – each word is modeled as if chosen at random from a specific topic by rolling a weighted w-sided die, where w is the total number of words in the meeting. Similarly, each utterance is represented as a multinomial distribution (θ) over topics. The parameters (i.e., the die-weights) for each multinomial distribution are themselves drawn from a uniform *Dirichlet* distribution with *hyperparameters* α, for topics, and β for words. The Dirichlet distribution is a multivariate probability distribution that is the conjugate prior to the multinomial distribution. Its hyperparameters control how broad or specific topics are as discussed below.

A. Implementing Bayesian Topic Models

Several implementations of LDA and related algorithms have been made available for public use (e.g., [81–83]). We have implemented one such algorithm in Python which is freely available online at http://code.google.com/p/knowledge-boundaries/. Such implementations fit the model to a specific corpus using Bayesian inference algorithms. In particular, we are interested in finding the most probable hypothesis, h, (i.e., the most appropriate model),

given the observed data, d (i.e., the meeting transcript). Model fitting may be explained using Bayes' theorem:

$$p(h|d) = \frac{p(d|h)^* p(h)}{p(d)} \qquad \text{(A1-1)}$$

or using the notation specific to the LDA model [66]:

$$p(\theta, z|w, \alpha, \beta) = \frac{p(\theta, z, w|\alpha, \beta)}{p(w|\alpha, \beta)} \qquad \text{(A1-2)}$$

This Bayesian formulation assures that the topics that are inferred by LDA are appropriate to the corpus being analyzed and that the model is not over-fit to the corpus data.

Explicit computation of LDA's posterior distribution (i.e., the distribution that we would like to determine in order to be able to fit topics to the data) is intractable. To see why this is, we must expand the expression above into its constituent parts. The numerator is easily expanded using the canonical expressions for the multinomial and Dirichlet distributions.

$$p(\theta, z, w|\alpha, \beta) = p(w|\phi)^* p(\phi|\beta)^* p(z|\theta)^* p(\theta|\alpha) = C_1(\alpha, \beta)^* (\prod_{i=1}^{V} \phi_i^{w_i})^* (\prod_{j=1}^{V} \phi_j^{\beta-1})^* (\prod_{k=1}^{T} \theta_t^{z_t})^* (\prod_{k=1}^{T} \theta_t^{\alpha-1})$$

$$\text{(A1-3)}$$

Here, V is the total number of words in the corpus and T is the total number of topics. C_1 is a constant term that depends only on the hyperparameters, and serves to normalize the distribution. We may interpret C_1 as presumed data that has already been seen and added to the observed data. The hyperparameters controlling the behavior of the Dirichlet distribution may therefore be said to reflect one's prior beliefs regarding the propensity of a particular topic or word in the data. As will be discussed below, these hyperparameters are set to standard, empirically-tested values. The denominator is not analytically tractable [66]:

$$p(w|\alpha, \beta) = C_2(\alpha)^* \int (\prod_{i=1}^{T} \theta_i^{\alpha-1})^* (\prod_{n=1}^{D} \prod_{j=1}^{T} \prod_{k=1}^{V} (\theta_j \beta_{jk})^{w_n^k}) d\theta \qquad \text{(A1-4)}$$

Consequently, the posterior distribution must be estimated, as will be described below.

B. The Author-Topic Model

The LDA model as outlined above is still sensitive to the arbitrary document boundaries imposed by the court recorder. Furthermore, documents (i.e., utterances) vary significantly in length – some might only be two words (e.g., "Thank you") whereas others might be

significant monologues. A variant of LDA, the Author-Topic (AT) model, can been used to generate a distribution over topics for each participant in a meeting [84] that is insensitive to document boundaries. Since the speaker's identity guides topic formation, shared topics are more likely to represent common jargon. The Author-Topic model provides an analysis that is guided by the authorship data of the documents in the corpus, in addition to the word co-occurrence data used by LDA. Each author (in this case, a speaker in the discourse), rather than each utterance, is modeled as a multinomial distribution over a fixed number of topics that is pre-set by the modeler. Each topic is modeled as a multinomial distribution over words.

C. Estimating the posterior distribution

Most popular implementations of LDA estimate the posterior distribution for its test data using Gibbs sampling – a Markov Chain Monte Carlo (MCMC) technique adopted from statistical physics (e.g., [16]).

Details of the MCMC algorithm derivation for the AT Model are given in [84]. The AT model was implemented in Python and MATLAB by the authors, based on [16]:

Initialize topic assignments randomly for all word instances

```
repeat
  for d=1 to D do
    for i=1 to N_d do
      draw z_di & x_di from P(x_di,z_di|z_−di, x_−di w,α,β)
      assign z_di & x_di and update count vectors
    end for
  end for
until Markov chain reaches equilibrium
```

Here, D is the total number of documents, N_d is the number of word tokens in each document, z_{di} is the topic of word token i in document d, and x_{di} is the author assigned to word token i in document d. The form of $P(x_{di},z_{di}|\mathbf{z}_{-di}, \mathbf{x}_{-di} \mathbf{w},\alpha,\beta)$ is derived in [84] as follows:

$$P(z_{\mathrm{di}}=j, x_{\mathrm{di}}=k|z_{-\mathrm{di}}, x_{-\mathrm{di}}w, \alpha, \beta) \propto \frac{n_{-\mathrm{di},j}^{(w_{\mathrm{di}})}+\beta}{n_{-\mathrm{di},j}^{(\cdot)}+V\beta} \frac{n_{-\mathrm{di},j}^{(x_{\mathrm{di}})}+\alpha}{n_{-\mathrm{di},j}^{(\cdot)}+T\alpha} \qquad \text{(A1-5)}$$

Each word's probability of assignment to a given topic is proportional to the number of times that that word appears in that same topic, and to the number of times a word from that author is assigned to that topic. This defines a Markov chain, whose probability of being in given state is guaranteed to converge to the posterior distribution of the AT model given the corpus data, after a sufficiently large number of iterations.

D. How many topics?

We calculated the cross entropy of topic models with 1–35 topics for each transcript with the smallest number of topics that minimized cross-entropy chosen as the final number of topics for each transcript (full details are provided in [65]). This approach yielded a number of topics between 10 and 28, depending on the specific meeting being analyzed.

E. Hyperparameter selection

Like LDA, the AT model requires the selection of two parameters. Ideally, we would like to determine which parameters used for the AT model best fit the corpus data. We must be wary of over-constraining the analysis with the assumptions underlying the AT model. A popular metric for goodness-of-fit used within the machine learning literature is *cross-entropy* [85]. Cross-entropy is a metric of the average number of bits required to describe the position of each word in the corpus and is closely related to *log-likelihood*, a measure of how well a given model predicts a given corpus. Therefore, lower cross-entropy indicates a more parsimonious model fit and that the assumptions underlying the model are descriptive of the data. For the AT model, cross-entropy may be calculated as follows:

$$H(p,q) = -\frac{\sum\limits_{i=1}^{N} \log_2(p(w_i|\phi,\beta)^* p(z_i|\theta,\alpha)^* p(x_i))}{N} \qquad \text{(A1-6)}$$

In this expression, N is the total number of word tokens. The expression in the numerator is the empirical *log-likelihood*. Thus, a natural interpretation of cross-entropy is the average log-likelihood across all observed word instances in the corpus. The lower a given model's cross-entropy, or the higher its log-likelihood, the more parsimonious is the model's fit to the data.

Each author's topic distribution is modeled as having been drawn from a symmetric Dirichlet distribution, with parameter α. Values of α that are smaller than one will tend to more closely fit the author-specific topic distribution to observed data – if α is too small, one runs the risk of overfitting the data. Similarly, values of α greater than one tend to bring author-specific topic distributions closer to uniformity. This can be advantageous if we do not the identity of the author to strongly influence topic assignment. A value of $\alpha=50/(\#$ topics) was used for the results presented in this paper, based upon the values suggested by [16]. For the numbers of topics considered in these analyses, this corresponds to a mild smoothing across authors. Similar to α is the second Dirichlet parameter, β, from which the topic-specific word distributions are drawn. β values that are large tend to induce very broad topics with much overlap, whereas smaller values of β induce topics which are specific to small numbers of words. Following the empirical guidelines set forth by [16], and empirical testing performed by the authors, we set the value of $\beta = 200/(\#$ words). Given that the average corpus generally consists of ~25,000 word tokens, representing about 2500 unique words in about 1200 utterances, the value of β is generally on the order of 0.1, a value close to that used by [84]. Thus, topics tend to be defined by a small number of words. As will be

shown below, values of α tend to be on the order of 1 – 5 suggesting that the identity of a given speaker does not overly constrain a topic.

F. Committee filtering

Our analysis primarily focuses on the voting members on an advisory panel. This decision was made because it is precisely these members whose evaluations will determine the panel recommendations. Other, non-voting, panel members are not included as part of the committee in the following analyses because they play a relatively small role in panel discussion in the meetings examined. Inclusion of these members is straightforward, and examination of their roles is left to future research.

Panel members share certain language in common including procedural words and domain-specific words that are sufficiently frequent as to prevent good topic identification. As a result, a large proportion of the words spoken by each committee member may be assigned to the same topic, preventing the AT model from identifying important differences between speakers. In a variant of a technique suggested by [86], [65] solved this problem using the AT model by creating a "false author" named "committee". Prior to running the AT model's algorithm, all committee voting members' statements are labeled with two possible authors – the actual speaker and "committee". Since the AT model's algorithm randomizes over all possible authors, words that are held in common to all committee members are assigned to "committee", whereas words that are unique to each speaker are assigned to that speaker. In practice, this allows individual committee members' unique topic profiles to be identified, as demonstrated below. In the unlikely case where all committee members' language is common, half of all words will be assigned to "committee" and the other half will be assigned at random to the individual speakers in such a way as to preserve the initial distribution of that author's words over topics. This filtering technique greatly improves the accuracy of the method and is uniquely feasible with a Bayesian approach.

References

1. Carlile PR. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundariesgh. Organization science. 2004; 15(5):555–568.

2. Dougherty D. Interpretive barriers to successful product innovation in large firms. Organization science. 1992; 3(2):179–202.

3. Leonard-Barton D. Wellsprings of knowledge: Building and sustaining the sources of innovation. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship. 1995

4. Lovelace K, Shapiro DL, Weingart LR. Maximizing cross-functional new product teams' innovativeness and constraint adherence: A conflict communications perspective. Academy of management journal. 2001; 44(4):779–793.

5. Carlile PR. A pragmatic view of knowledge and boundaries: Boundary objects in new product development. Organization science. 2002; 13(4):442–455.

6. Bechky BA. Sharing meaning across occupational communities: The transformation of understanding on a production floor. Organization science. 2003; 14(3):312–330.

7. Boland RJ Jr, Tenkasi RV. Perspective making and perspective taking in communities of knowing. Organization science. 1995; 6(4):350–372.

8. Cramton CD. The mutual knowledge problem and its consequences for dispersed collaboration. Organization science. 2001; 12(3):346–371.

9. Majchrzak A, More PH, Faraj S. Transcending knowledge differences in cross-functional teams. Organization Science. 2012; 23(4):951–970.

10. Sole D, Edmondson A. Situated knowledge and learning in dispersed teams. British journal of management. 2002; 13(S2):S17–S34.

11. Bechky BA. Object Lessons: Workplace Artifacts as Representations of Occupational Jurisdiction. American Journal of Sociology. 2003; 109(3):720–752.

12. Kane AA. Unlocking knowledge transfer potential: Knowledge demonstrability and superordinate social identity. Organization Science. 2010; 21(3):643–660.

13. Star SL, Griesemer JR. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. Social studies of science. 1989; 19(3):387–420.

14. Tsoukas H. A dialogical approach to the creation of new knowledge in organizations. Organization Science. 2009; 20(6):941–957.

15. Mørk BE, Hoholm T, Maaninen-Olsson E, Aanestad M. Changing practice through boundary organizing: A case from medical R&D. Human Relations. 2012; 65(2):263–288.

16. Griffiths TL, Steyvers M. Finding scientific topics. Proceedings of the National Academy of Sciences. 2004; 101(suppl 1):5228–5235.

17. Orlikowski WJ. Sociomaterial practices: Exploring technology at work. Organization studies. 2007; 28(9):1435–1448.

18. Barnett WP, Hansen MT. The red queen in organizational evolution. Strategic Management Journal. 1996; 17(S1):139–157.

19. Polanyi M. Tacit knowing: Its bearing on some problems of philosophy. Reviews of modern physics. 1962; 34(4):601.

20. Weaver SJ, Rosen MA, DiazGranados D, Lazzara EH, Lyons R, Salas E, Knych SA, McKeever M, Adler L, Barker M, et al. Does teamwork improve performance in the operating room? A multilevel evaluation. The Joint Commission Journal on Quality and Patient Safety. 2010; 36(3): 133–142. [PubMed: 20235415]

21. Weisberg RW. Expertise and Reason in Creative Thinking: Evidence from Case Studies and the Laboratory. 2006

22. Chase WG, Simon HA. Perception in chess. Cognitive psychology. 1973; 4(1):55–81.

23. Faraj S, Sproull L. Coordinating expertise in software development teams. Management science. 2000; 46(12):1554–1568.

24. Gigerenzer, G. Adaptive thinking: Rationality in the real world. Oxford: University Press, USA; 2000.

25. Gigerenzer, G., Selten, R. Bounded rationality: The adaptive toolbox. Mit Press; 2002.

26. Reyna VF, Lloyd FJ. Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. Journal of Experimental Psychology: Applied. 2006; 12(3):179. [PubMed: 16953744]

27. Duhaime IM, Schwenk CR. Conjectures on cognitive simplification in acquisition and divestment decision making. Academy of Management Review. 1985; 10(2):287–295.

28. Gary MS, Wood RE, Pillinger T. Enhancing mental models, analogical transfer, and performance in strategic decision making. Strategic Management Journal. 2012; 33(11):1229–1246.

29. Gavetti G, Levinthal DA, Rivkin JW. Strategy making in novel and complex worlds: the power of analogy. Strategic Management Journal. 2005; 26(8):691–712.

30. Salas E, Rosen MA, Burke CS, Goodwin GF, Fiore SM. The Making of a Dream Team: When Expert Teams Do Best. 2006

31. Schön, DA. The reflective practitioner: How professionals think in action. Vol. 5126. Basic books; 1983.

32. Schön DA. Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. San Francisco. 1987

33. Szulanski G. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Strategic management journal. 1996; 17(S2):27–43.

34. Szulanski G. The process of knowledge transfer: A diachronic analysis of stickiness. Organizational behavior and human decision processes. 2000; 82(1):9–27.

35. Nonaka I, Toyama R, Konno N. SECI, Ba and leadership: a unified model of dynamic knowledge creation. Long range planning. 2000; 33(1):5–34.

36. Cook SD, Brown JS. Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. Organization science. 1999; 10(4):381–400.

37. Hargadon AB, Bechky BA. When collections of creatives become creative collectives: A field study of problem solving at work. Organization Science. 2006; 17(4):484–500.

38. Nonaka I. A dynamic theory of organizational knowledge creation. Organization science. 1994; 5(1):14–37.

39. Shannon CE, Weaver W. The mathematical theory of information. 1949

40. Dougherty D. Reimagining the Differentiation and Integration of Work for Sustained Product Innovation. Organization Science. Oct; 2001 12(5):612–631.

41. Edmondson AC, Nembhard IM. Product development and learning in project teams: the challenges are the benefits. Journal of Product Innovation Management. 2009; 26(2):123–138.

42. Hackman, JR. Leading teams: Setting the stage for great performances. Harvard: Business Press; 2002.

43. Nonaka, I., Takeuchi, H. The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford: university press; 1995.

44. Van Der Vegt GS, Bunderson JS. Learning and performance in multidisciplinary teams: The importance of collective team identification. Academy of Management Journal. 2005; 48(3):532–547.

45. Akgün AE, Byrne JC, Lynn GS, Keskin H. Organizational unlearning as changes in beliefs and routines in organizations. Journal of Organizational Change Management. 2007; 20(6):794–812.

46. Klein G, Phillips JK, Rall EL, Peluso DA. A data-frame theory of sensemaking. Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making. 2007:113–155.

47. Weick, KE. Sensemaking in organizations. Vol. 3. Sage: 1995.

48. Weick KE, Sutcliffe KM, Obstfeld D. Organizing and the process of sensemaking. Organization science. 2005; 16(4):409–421.

49. Endsley MR. Toward a theory of situation awareness in dynamic systems. Human Factors: The Journal of the Human Factors and Ergonomics Society. 1995; 37(1):32–64.

50. Mosier KL. Expert decision-making strategies. 1991

51. Faraj S, Xiao Y. Coordination in fast-response organizations. Management science. 2006; 52(8):1155–1169.

52. Lamont, M. How professors think: Inside the curious world of academic judgment. Harvard: University Press; 2009.

53. Stasser G, Titus W. Pooling of unshared information in group decision making: Biased information sampling during discussion. Journal of personality and social psychology. 1985; 48(6):1467.

54. Stasser G, Stewart DD, Wittenbaum GM. Expert roles and information exchange during discussion: The importance of knowing who knows what. Journal of experimental social psychology. 1995; 31(3):244–265.

55. Stasser G, Titus W. Hidden profiles: A brief history. Psychological Inquiry. 2003; 14(3–4):304–313.

56. Burt, RS. Structural holes: The social structure of competition. Harvard: university press; 2009.

57. Tortoriello M, Krackhardt D. Activating cross-boundary knowledge: The role of Simmelian ties in the generation of innovations. Academy of Management Journal. 2010; 53(1):167–181.

58. Reagans R, McEvily B. Network structure and knowledge transfer: The effects of cohesion and range. Administrative science quarterly. 2003; 48(2):240–267.

59. Tortoriello M, Reagans R, McEvily B. Bridging the knowledge gap: The influence of strong ties, network cohesion, and network range on the transfer of knowledge between organizational units. Organization Science. 2012; 23(4):1024–1039.

60. Argote L, Ingram P. Knowledge transfer: A basis for competitive advantage in firms. Organizational behavior and human decision processes. 2000; 82(1):150–169.

61. Brown JS, Duguid P. Knowledge and organization: A social-practice perspective. Organization science. 2001; 12(2):198–213.

62. Phelps C, Heidl R, Wadhwa A. Knowledge, networks, and knowledge networks a review and research agenda. Journal of Management. 2012; 38(4):1115–1166.

63. Wasserman, S., Faust, K. Social network analysis: Methods and applications. Vol. 8. Cambridge: university press; 1994.

64. Griffiths TL, Steyvers M, Tenenbaum JB. Topics in semantic representation. Psychological review. 2007; 114(2):211. [PubMed: 17500626]

65. Broniatowski DA, Magee CL. Studying Group Behaviors: A tutorial on text and network analysis methods. Signal Processing Magazine, IEEE. 2012; 29(2):22–32.

66. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. the Journal of machine Learning research. 2003; 3:993–1022.

67. Blei DM. Probabilistic topic models. Communications of the ACM. 2012; 55(4):77–84.

68. Quinn KM, Monroe BL, Colaresi M, Crespin MH, Radev DR. How to analyze political attention with minimal assumptions and costs. American Journal of Political Science. 2010; 54(1):209–228.

69. [Accessed: 21-Sep-2016] CDRH Advisory Meeting Materials Archive. [Online]. Available: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfAdvisory/details.cfm?mtg=150

70. Bavry AA, Kumbhani DJ, Helton TJ, Borek PP, Mood GR, Bhatt DL. Late thrombosis of drug-eluting stents: a meta-analysis of randomized clinical trials. The American journal of medicine. 2006; 119(12):1056–1061. [PubMed: 17145250]

71. Camenzind E. Treatment of In-Stent Restenosis-Back to the Future? New England Journal of Medicine. 2006; 355(20):2149. [PubMed: 17101616]

72. Nordmann AJ, Briel M, Bucher HC. Mortality in randomized controlled trials comparing drug-eluting vs. bare metal stents in coronary artery disease: a meta-analysis. European heart journal. 2006; 27(23):2784–2814. [PubMed: 17020889]

73. Rittel HW, Webber MM. Dilemmas in a general theory of planning. Policy sciences. 1973; 4(2): 155–169.

74. Ackoff RL. The future of operational research is past. Journal of the operational research society. 1979; 30(2):93–104.

75. Janis IL. Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes. 1972

76. Franco LA. Rethinking soft OR interventions: models as boundary objects. European Journal of Operational Research. 2013; 231(3):720–733.

77. Pidd M, Woolley RN. A pilot study of problem structuring. Journal of the operational research society. 1980; 31(12):1063–1068.

78. Dong A. The latent semantic approach to studying design team communication. Design Studies. 2005; 26(5):445–461.

80. De Finetti B. Theory of Probability. A critical introductory treatment. 1979

81. McCallum AK. Mallet: A machine learning for language toolkit. 2002

82. Manning CD, Surdeanu M, Bauer J, Finkel JR, Bethard S, McClosky D. The Stanford CoreNLP Natural Language Processing Toolkit. ACL (System Demonstrations). 2014:55–60.

83. Steyvers, M., Griffiths, T. Matlab topic modeling toolbox 1.4. 2011. URL http://psiexp.ss.uci.edu/research/programs_data/toolbox.htm

84. Rosen-Zvi, M., Griffiths, T., Steyvers, M., Smyth, P. The author-topic model for authors and documents; Proceedings of the 20th conference on Uncertainty in artificial intelligence; 2004. p. 487-494.

85. Manning, CD., Schütze, H. Foundations of statistical natural language processing. Vol. 999. MIT Press; 1999.

86. Rosen-Zvi M, Chemudugunta C, Griffiths T, Smyth P, Steyvers M. Learning author-topic models from text corpora. ACM Transactions on Information Systems (TOIS). 2010; 28(1):4.

Biographies

David A. Broniatowski (M'10) was born in Cleveland, Ohio, in 1982. He received an S.B degree in aero/astro engineering in 2004, S.M degrees in aero/astro engineering and Technology and Policy, and a Ph.D. degree in Engineering Systems with a focus on Technology Management and Policy in 2010, all from the Massachusetts Institute of Technology.

From 2010 to 2012, he was Senior Research Scientist at Synexxus, Inc., a systems integration firm, and from 2012–2013, he was a Postdoctoral Fellow at the Center for Advanced Modeling in the Social, Behavioral, and Health Sciences in the Johns Hopkins University School of Medicine's Department of Emergency Medicine. Since 2013, he has been an Assistant Professor with the Department of Engineering Management and Systems Engineering at The George Washington University. His work has appeared in journals including *Science, IEEE Signal Processing Magazine, Systems Engineering, the American Journal of Preventive Medicine, Medical Decision Making, the American Journal of Therapeutics, Vaccine*, and others. He has received research support from the National Institutes of Health.

Christopher L. Magee received the B.S., M.S., and Ph.D. degrees in metallurgy and materials science from the Carnegie Institute of Technology (now Carnegie Mellon University), Pittsburgh, PA, and the M.B.A. degree from Michigan State University, East Lansing.

He is currently Professor of the Practice, Institute for Data, Systems and Society (IDSS) and the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, and the Co-Director of the SUTD/MIT International Design Center. His recent research has focused upon empirical and theoretical quantification of technological progress as well as innovation in complex systems.

Dr. Magee was elected to the National Academy of Engineering while with Ford Motor Company for contributions in advanced vehicle development in 1997. He was a Ford Technical Fellow in 1996 and is a Fellow of the American Society for Materials.
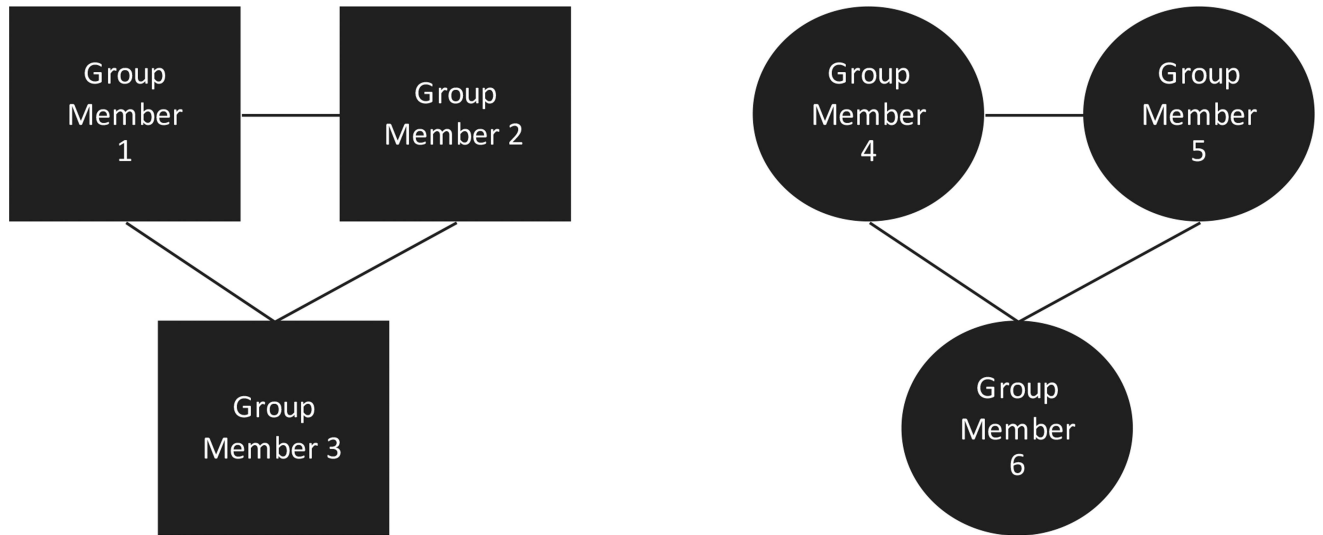
Fig. 1.
An idealization of a semantically cohesive network, reflecting the existence of a semantic boundary. Members of the same community of practice (sharing the same shape) are linked, but there are no connections between communities of practice.
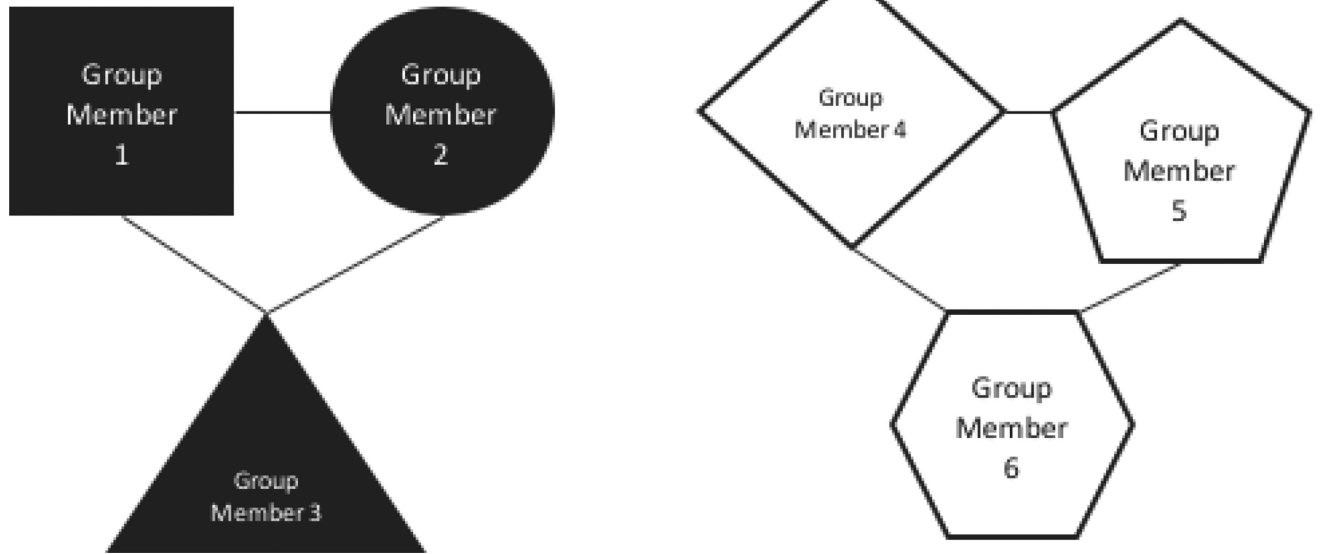
Fig. 2.
An idealization of a network reflecting the existence of a pragmatic boundary. Members who vote the same way (sharing the same color) are linked. Again, different shapes represent different communities of practice.
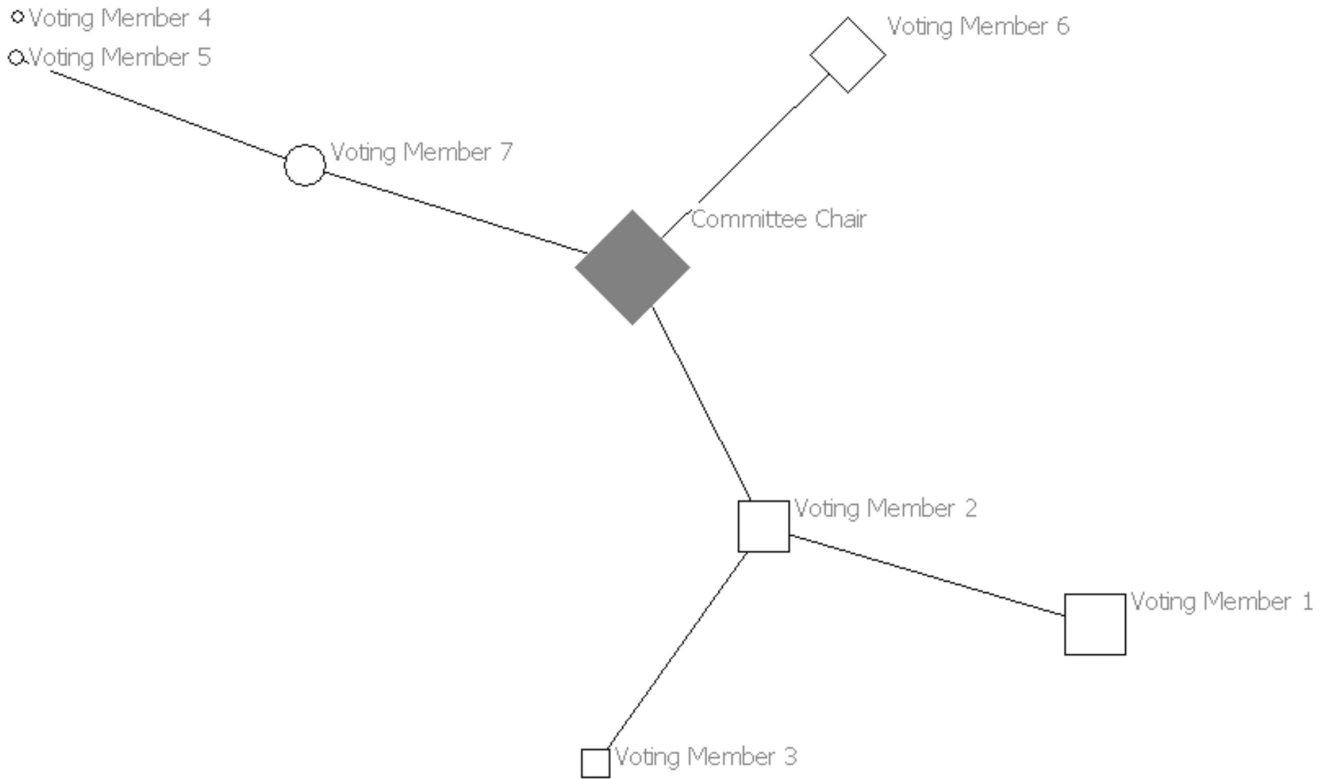
Fig. 3.
Graph of the FDA Circulatory Systems Advisory Panel meeting held on December 5, 2000. Node shape represents medical specialty (squares are cardiologists; diamonds are electrophysiologists; circles are surgeons). Node size is proportional to the number of words spoken by that group member during this meeting. All participants voted for device approval. The committee chair is in grey.
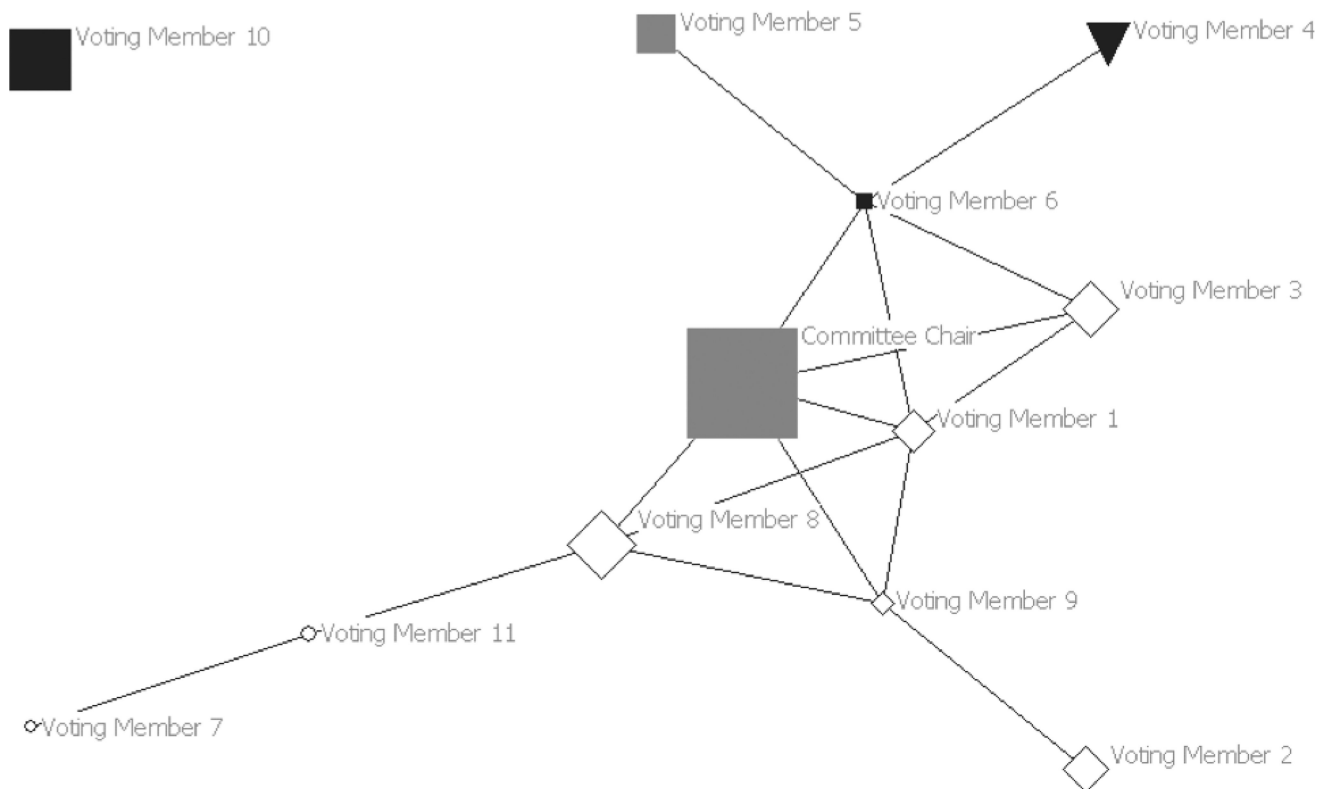
Fig. 4.
Graph of the FDA Circulatory Systems Advisory Panel meeting held on March 6, 2003. Node shape represents medical specialty (squares are cardiologists; diamonds are electrophysiologists; circles are surgeons; the triangle is a statistician). White nodes voted in favor of device approval; black nodes voted against device approval; grey nodes did not vote. Node size is proportional to the number of words spoken by that group member during this meeting. The committee chair is in grey.
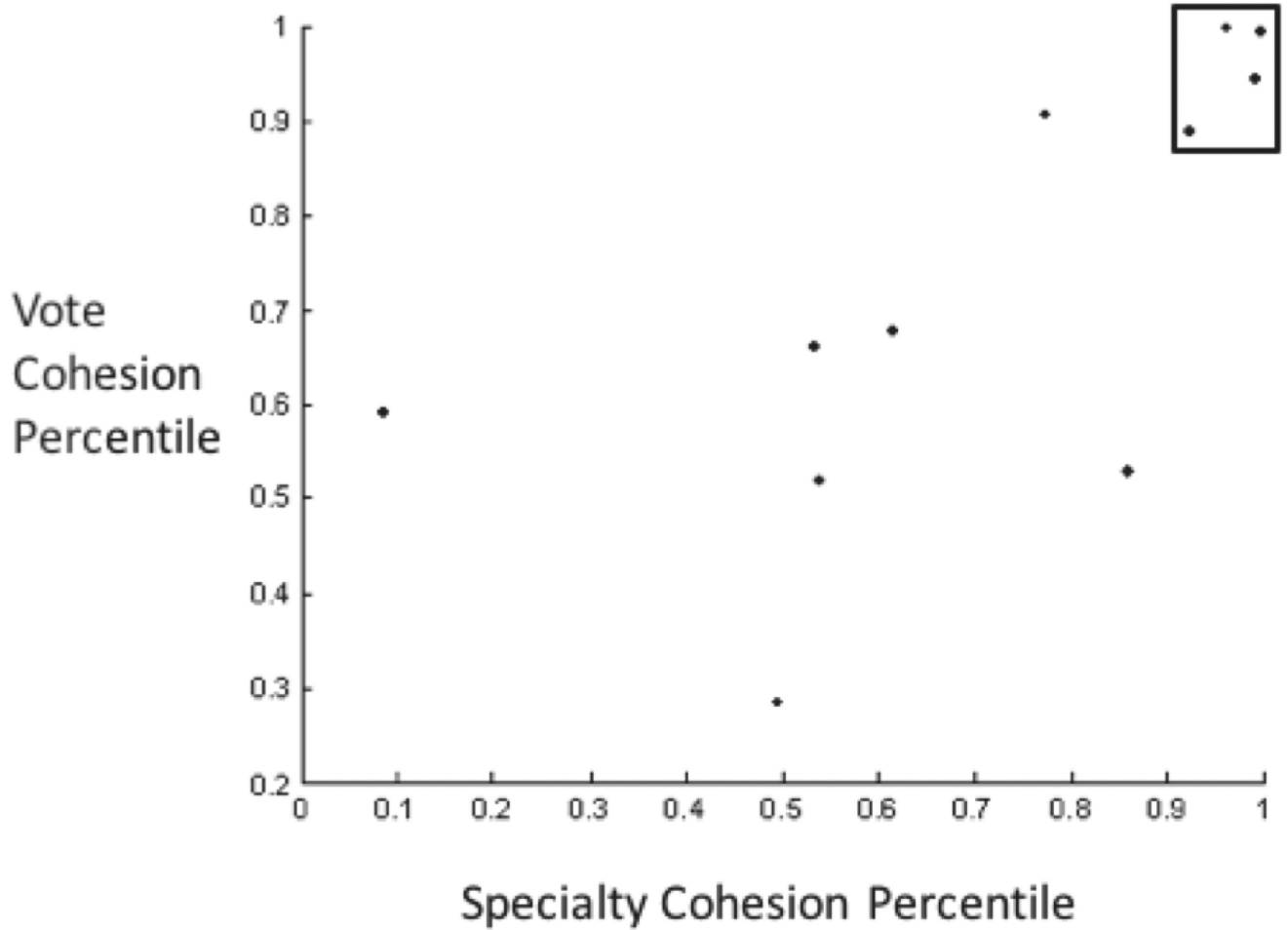
Fig. 5.
Scatter plot of Vote Cohesion percentile vs. Specialty Cohesion percentile for 11 meetings in which there was a minority of two or more. Vote and specialty cohesion percentiles are positively associated (Spearman Rho =0.79; p=0.006). Note especially the four points clustered in the upper-right of the graph (boxed) which demonstrate that high vote cohesion occurred only when specialty cohesion is high.

TABLE I

The top five word-stems for one run of the AT model on the corpus for the Circulatory Systems Devices Panel Meeting of March 4, 2002.

Topic	Word Stems				
1	clinic	endpoint	efficaci	comment	base
2	trial	insync	icd	studi	was
3	was	were	sponsor	just	question
4	patient	heart	group	were	failure
5	devic	panel	pleas	approv	recommend
6	think	would	patient	question	don
7	dr	condit	vote	data	panel
8	effect	just	trail	look	would
9	lead	implant	complic	ventricular	event
10	patient	pace	lead	were	devic

TABLE II

Summary of qualitative analysis for seven meetings. These meetings were chosen as representative of the range of novelty encountered by the Circulatory Systems Devices Advisory Panel

Meeting Number	SCP	VCP	Novelty
1	0.968	N/A	Low
2	0.975	N/A	Low
3	0.959	1.00	Moderate
4	0.996	0.987	Moderate
5	0.087	0.586	High
6	0.494	0.277	High
7	0.362	N/A	High

SCP = Specialty Cohesion Percentile; VCP = Vote Cohesion Percentile; N/A = Not Applicable because of voting minority less than 2.