



Full length article

Advice reification, learning, and emergent collective intelligence in online health support communities

Joshua Introne^{a,*}, Sean Goggins^b^a Department of Media and Information, Michigan State University, 404 Wilson Rd., East Lansing, MI, 48824, USA^b Electrical Engineering and Computer Science, University of Missouri, W1024 Lafferre Hall, Columbia MO, 65211, USA

ARTICLE INFO

Keywords:

Online health community
Health advice
Community of practice
Reification

ABSTRACT

Online health support forums utilize straightforward online discussion designs to create a sociotechnical space where people can seek social support from others. The advice generated in these forums exists as an archival resource for future health information seekers. The present study uses mixed methods to investigate how invisible social processes lead advice to be adapted over time by forum members. Drawing on the construct of 'reification' from the communities of practice (COP) literature, we operationalize the reification of advice (RoA) as a process by which advice is developed across multiple discussion threads, and construct an algorithmic procedure to extract posts that trace this process. We evaluate our algorithm with crowd-workers, and perform an inductive, qualitative analysis to identify different modes of advice reification. We suggest that RoA could be used as the basis of a mid-level theory that treats online support communities and bundles of advice trajectories embedded in a shifting sociotechnical context. In our closing analysis, we propose that our approach might be a first step in an algorithmic procedure for assessing advice quality, drawing on the idea that reified advice may be considered a product of the collective intelligence of an online health support community.

1. Introduction

People who participate in online social platforms leave digital traces of their activities, which can in turn become a resource for others (e.g. articles on Wikipedia, code on GitHub, answers on StackOverflow). Social computing platforms employ sociotechnical infrastructures designed to scaffold longitudinal process that adapt and, hopefully, improve the quality of these resources (Bryant, Forte, & Bruckman, 2005; Crowston, Wei, Li, & Howison, 2006; Parnin, Treude, & Grammel, 2012; Solomon & Wash, 2014). Highlighting their ability to harness the collective efforts of online contributors, such social computing platforms have been described as a kind of designed collective intelligence (Kittur & Kraut, 2008; Malone, Laubacher, & Dellarocas, 2009; Quinn & Bederson, 2011).

In other cases, resources generated by online platforms are a side effect of online engagement. Such is often the case with online health support communities (OHC). OHCs are places where people can go to find social support for health conditions. In many cases OHCs are organized as Question & Answer style discussions, and built upon well-established threaded discussion forum technologies (Davison, Pennebaker, & Dickerson, 2000; Maloney-Krichmar & Preece, 2005). The first post in each threaded conversation is usually a question (a

request for some kind of support) and the remainder of the conversation is devoted in part to responding to the initial question while also providing long-time forum members with an opportunity for richer interactions (Introne, Semaan, & Goggins, 2016). These conversations often yield advice for the original asker, and this advice may also be considered a resource for future seekers.

Unlike social computing platforms, the processes that generate advice in an OHC are not governed by any designed sociotechnical infrastructures. However, research has shown that in some OHCs, the same small, tightly knit group of people—the community's *core*—responds to thousands of requests from a much larger population of support seekers (Bambina, 2007; Introne et al., 2016). Given this, it seems likely that advice stabilizes as these tightly knit groups of people revisit the same kinds of repeated support requests. Unlike the designed collective intelligence of other social computing platforms, the stabilization and refinement of advice suggests a form of emergent collective intelligence in OHCs.

The primary aim of this paper is to develop evidence for the existence of such emergent collective intelligence. To do this, we draw heavily on Wenger's "communities of practice" (COP) construct (Lave & Wenger, 1991; Wenger, 1998), and the concomitant process of *reification* that occurs therein. A COP is a group of people who share a craft or

* Corresponding author.

E-mail addresses: jintrone@msu.edu (J. Introne), GogginsS@missouri.edu (S. Goggins).

profession, and its participants reify their practice-oriented experiences over time. Wenger (1998) defines the term reification thusly:

“With the term reification I mean to cover a wide range of processes that include making, designing, representing, naming, encoding and describing, as well as perceiving, interpreting, using, reusing, decoding and recasting. Reification occupies much of our collective energy: from entries in a journal to historical records, from poems to encyclopaedias, from names to classification systems, from dolmens to space probes ... In all these cases, aspects of human experience and practice are congealed into fixed forms and given the status of object.”

Although reification may produce a concrete object, it is not primarily a process of objectification. Rather, it is a process through which aspects of participation in a community of practice are made more concrete, enabling these aspects to become foci for further negotiation (Wenger, 1998, p. 58). We apply this notion of reification to describe the processes by which members in an OHC produce and adapt health advice. Health advice in an OHC is a complex object that emerges from individuals' shared practice of living with a health condition, and includes personal experiences with that condition and the healthcare system, sensibilities about how to guide others, and collective sense-making around these things both on and offline. We argue that the advice produced online reifies this amalgam of experiences via the active participation of individuals grappling with a similar health condition in the virtual space of an OHC.

Adopting this lens, we use this paper first to identify and then classify instances of advice reification across multiple forums hosted by the population an online health information service, WebMD. Our data is drawn from five years across fifty-five different, health-condition specific forums. To perform our analysis, we introduce a novel algorithmic approach to extracting sequences of posts—which we call *meta-conversations (MCs)*—that are arrayed across discussion threads and are likely vehicles for advice reification. We use a content analytic procedure to validate that these sequences of posts are indeed likely to trace advice reification processes. Then, we use qualitative methods to examine the structure and content of these MCs in order to isolate specific genres (Andersen & van Leeuwen, 2017; Caple & Knox, 2017) of advice reification. Finally, in our closing discussion, we consider both the theoretical implications of our findings as well as the practical application of our techniques to the problem of assessing advice quality online.

1.1. Background

Online health support platforms are online spaces where people go to find some form of social support from others (Davison et al., 2000; Maloney-Krichmar & Preece, 2005). Access to social support can have a range of clinical benefits for people with health conditions, and in particular those with chronic conditions (Fox & Purcell, 2010). It is thought that because online support reduces barriers to access and can insulate people from stigma, that it may be a valuable tool for many different patient populations (Davison et al., 2000; Johnson & Ambrose, 2006). An important component of online social support involves advice about how to manage a health condition.

A pressing concern for medical professionals and researchers is whether or not advice found online is of high quality. The existence of misleading or inaccurate information about health issues on the web (though not necessarily within online social support platforms) has been well documented. For example, a systematic review of online health communities found that 55 of 79 distinct studies found information quality to be a problem (Eysenbach, Powell, Kuss, & Sa, 2002). However, online advice is not uniformly bad, and evidence suggests that different communities may develop practices that influence its quality (Hartzler & Huh, 2016). Indeed, several studies have found that in some forums, members actively monitor for inaccuracies

and correct ‘bad’ advice (Deshpande & Jadad, 2006; Esquivel, Meric-Bernstam, & Bernstam, 2006).

An important, but as yet underdeveloped line of research is how of social structures and practices that develop in OHCs bear upon the production of advice. Members in OHCs can come to play social roles that interact with advice production in more or less direct ways. For example, Maloney-Krichmar and Preece (2005) identified 17 social roles in a support group for people with knee problems, and a subset of these were explicitly focused on advice and its quality (e.g. “evaluator-critics” and “information givers”). More generally, OHCs often stratify into core and more peripheral users (Bambina, 2007; Introne et al., 2016) who engage in different kinds of behaviors. Core members typically develop strong and convivial relationships with one another. Although these core groups are quite small—Introne et al. (2016) reports core sizes of tens of members or less—they can form persistent groups in OHCs that provide the bulk of advice to thousands of more peripheral members and newcomers (Introne et al., 2016).

A similar stratification from central to more peripheral participants can be found in communities of practice (COP) (Brown & Duguid, 2001; Lave & Wenger, 1991; Wenger, 1998). A COP may be considered to be a community that is characterized by its participants' mutual interest and participation in a common practice (Silva, Goel, & Mousavidin, 2009; Wenger, 1998, p. 72). As described by (Lave & Wenger, 1991), newcomers in a COP accumulate experience by working alongside more experienced members (a process referred to as legitimate peripheral participation) possibly becoming masters themselves who will guide subsequent newcomers. This process transforms the community's practice-based repertoire of knowledge, routines, and artifacts as newcomers potentially bring diverse perspectives and older members integrate, curate, and disseminate their accumulated repertoire.

Several researchers have called out both the differences and similarities between offline COPs and OHCs. Jones and Preece (2006) introduced the term “Community of Interest” to highlight the fact that the locus of an OHC is a common health concern, rather than a specific practice. Similarly, Johnston et al. (Johnston, Worrell, Di Gangi, & Wasko, 2013) distinguish OHCs from COPs because they lack a clear professional jargon. On the other hand, Kimmerle et al. (2013) focused on aspects of identity and knowledge management to argue that one online alternative health community could be seen as an excellent fit for a COP.

Unlike Jones and Preece (2006), we take the position that managing a health condition is indeed a central practice shared by members of at least some OHCs, and this is a critical part of what binds the people who find one another in these communities. Managing diabetes, or breast cancer, or fibromyalgia, or multiple sclerosis whilst trying to live a fulfilling life is not a dispassionate interest for these individuals. What and when to eat, how to navigate networks of medical specialists, and grappling with the emotional tumult of a debilitating condition all require the development of a complex web of strategies and behaviors, and sharing these with others who face the same challenges can be an important vehicle for learning and improvement.

We consider advice to be a concrete instantiation of the practice of managing a health condition. We anticipate that the core members of an OHC adapt advice over time for three reasons. First, an OHC is a public space, so the production of advice is visible and thereby becomes a potential target for deliberative engagement. Second, thousands of incoming requests from more peripheral individuals create opportunities for the core to produce the same kinds of advice time after time. Finally, core members are continuously learning about how to manage their own health conditions, and questions or offerings from newcomers may bring new information for consideration.

The concretization and re-concretization of the practice of condition management through the vehicle of advice-giving fits well with Wenger's construct of reification. We briefly discuss reification and its potential relationship to advice in OHCs below.

1.2. Reification and advice

Reification and participation are the two central processes that lie at the core of Wenger's COP framework (Wenger, 1998). Wenger considers these two processes as a duality rather than a dichotomy; they are intertwined, synergistic, and both fundamental to how COPs negotiate meaning and learn. Participation is the more easily defined of the two, and Wenger's usage is consonant with its dictionary definition: "To have or take a part or share with others (in some activity, enterprise, etc.)" (Wenger, 1998, p. 55). In the context of a COP, the activity or enterprise is social, and stems from membership in social communities.

Reification is more slippery, and in Wenger's usage is a process through which aspects of practice 'congeal' to something more stable. Some literature conflates the output of reification with the production of physical artifacts, but Wenger's definition is more nuanced:

Reification can take a great variety of forms: a fleeting smoke signal, or an age-old pyramid, an abstract formula or a concrete truck... a telling glance or a long silence, a private knot on a handkerchief or a controversial statue on a public square... What is important about all these objects is that they are only the tip of an iceberg, which indicates larger contexts of significance realized in human practices... Properly speaking, the products of reification are not simply concrete, material objects. Rather they are reflections of these practices, tokens of vast expanses of human meanings. (Wenger, 1998, pp. 60–61)

Two additional aspects of reification are important for our application of the concept to online advice. First, Wenger uses the term reification to refer to both the object of reification and the process that produces it, explaining that "if meaning exists only in its negotiation, at the level of meaning, the process and the product are not distinct" (Wenger, 1998, p. 60). Thus, we will use the term reification to refer to both advice and the processes that produce it. Second, while reification may "freeze" an aspect of practice, it does not also freeze meaning, any more than the meaning of the U.S. Constitution is frozen by the document itself. Rather, the reified object can help focus the negotiation of meaning for a COP, which is itself an ongoing process. That is, "the process of reification thus compels us to renegotiate the meaning of its past products, in the same way that a scar keeps bringing a past foolishness or heroic deed into conversations" (Wenger, 1998, p. 88).

Several analyses have hinted at this ongoing negotiation around advice in OHCs. For example, Huh and Ackerman (2012) describe how members in an online diabetes support community engage in a kind of collective sensemaking, through which they continuously refine their understanding of diabetes and how to manage it. In one example, forum members share a "startup solution kit," containing various resources and simple, easy to follow strategies that had been "well-polished through repetitive use" (Huh & Ackerman, 2012, p. 856). In a similar vein, Mamkykina et al. (2015) suggest that reification occurs in part via lateral engagement among users of a diabetes forum, which can lead to a transformation of ideas. Here, lateral engagement refers to the exchanges among heterogeneous users with different perspectives in long discussion threads.

Centering our analysis on the reification of advice in an online, discussion-based communities hold particular interest because of the particular form advice takes. The articulation of advice in a post is at once a reification, but this permanent artifact is itself somewhat ephemeral for the persistent, active members of the forum. While it may be recovered (e.g., via a search or bookmark), it is not ensconced in some highly salient artifact like a wiki article, and the focal point of member participation will move on to new conversations. Yet similar advice may be reified the next time a new seeker posts a similar request, affording an opportunity for adaptation. It is across a sequence of such reifications that we might observe the continuous negotiation of meaning, and perhaps learning. Our aim here to isolate and examine such longitudinal sequences. In the following sections, we outline our

approach.

1.3. Operationalizing reification

To trace the longitudinal reification of advice (RoA), we sought to identify sequential instances of the same advice embedded in posts that are arrayed across multiple threads. Prior studies (e.g., Huh & Ackerman, 2012; Mamkykina, Nakikj, & Elhadad, 2015) have examined individual discussion threads for related instances of collective information processing, but this constrains the time-scale of the processes these studies reveal. An important aim of our work is to elucidate sequences of RoA and the potential renegotiation of meaning these sequences imply across longer periods of time.

The primary virtual site of participation in most OHCs is the thread-based, asynchronous conversation. Conversation threads are usually initiated with a request for some kind of support, followed by a series of replies from multiple (possibly many) individuals. As described above, we anticipate that repeated encounters by the same members in the context of similar advice requests affords an opportunity for a continuous negotiation of meaning around advice. To operationalize this idea, we focus on two characteristics of this process—first, the posts that transcribe repeated instances of reification of the same advice will be roughly "about" the same thing, though the content may drift as advice is refined; second, these posts will proceed in sequence across, rather than within threads, and will be authored by individuals that have jointly participated in other conversations. We refer to such a set of posts as *meta-conversations* (MCs). Fig. 1 offers a schematic illustration of how a pair of posts in an MC is arranged across threads.

Note that we do not expect MCs to follow the organization of typical conversation; e.g., we do not expect turn-taking to occur (Sacks, Schegloff, & Jefferson, 1974), or for references to be resolved or maintained across posts (H. Clark & Marshall, 1981). Instead, we intend the term MC to describe something less than a conversation to coordinate understanding about a specific topic, but more than the epidemiological diffusion of a viral video.

More formally, an MC is a set of posts, arranged as a directed acyclic graph. Each post has four features: the time (t) at which it occurred, the thread in which it appears (T), the identity of the poster (u), and the topic (o) of the post. For us, a topic is an algorithmically derived indicator of what a post is "about," and hence a basis for an assumption about the similarity content across posts. A single link in the graph can be inferred from a tuple consisting of three posts:

$$[P_1(T_i, u_i, o_i), P_2(T_i, u_j, *), P_3(T_j, u_j, o_i)]$$

where $i \neq j$ for all subscripts, and subject to the following constraints:

$$t_1 < t_2 < t_3; t_3 - t_2 < k$$

$$Sim(P_{t_1}, P_{t_2}) < s$$

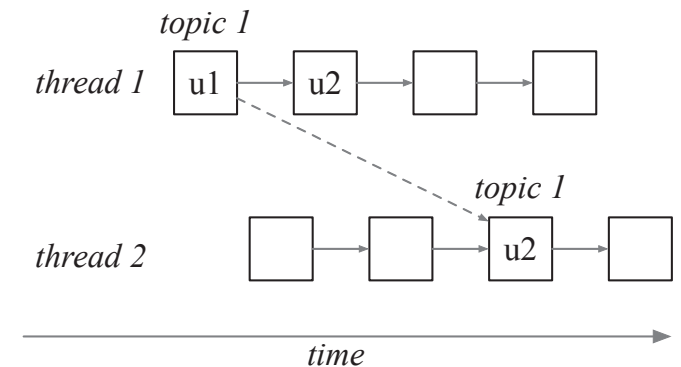


Fig. 1. Schematic for a single link (the dashed line) in a meta-conversation. Exposure of poster u2 to topic1 (posted by u1) results in u2 carrying topic1 to a new thread.

The constraints k and s capture two distinct intuitions about the nature of an MC. The constraint k indicates that after some period of time, we presume u_j has either forgotten about the original post, or is not part of an MC. The constraint s is an additional measure of lexical similarity, reflecting the idea that algorithmically derived topics may be too imprecise a classification to establish the joint membership of posts in a single MC.

A link in an MC can be thought of as a possible transfer of knowledge about reified advice between two users. The potential (observed) transfer occurs between the first and last post in the preceding tuple; the second post merely serves as evidence that the two posters (u_i and u_j) shared a common context and is not part of the meta-conversation. That is, there is a potential for transfer any time two people (u_i and u_j) participate (as evidenced by their posting activity) in a single thread, and then the latter person subsequently (u_j) posts in a different thread about the topic (o_i) originally posted by the first poster (u_i). In the methods section, we describe how we applied the above model to our dataset, and in particular, how we inferred values for s (lexical similarity between posts) and k (maximum time distance between two posts in different threads with a reasonable probability of those posts constituting a meta-conversation).

2. Material and methods

2.1. Corpus description

We drew our data from discussion communities hosted by WebMD, a relatively popular and long-standing site for health information. Several prior studies have examined these specific support communities (Huh, 2015; Introne et al., 2016; Kanthawala, Vermeesch, Given, & Huh, 2016; Ridings & Wasko, 2010), and our work builds on that described by Introne et al. (2016) in that it adopts a characterization of *coreness*; we describe users as being more core like as we traverse a scale from extra-periphery (individuals who post a handful of times and leave), to periphery (less active members with few observable online relationships), to core (highly active members with strong observable relationships). Conversations on WebMD are publicly viewable and do not contain identifying information. After consulting with the lead author's IRB, the institution did not consider this study to involve Human Subjects data because there was no manipulation of our subject population, and the data is publicly available. We include some quotes from the posts on the forum, but we have made sure that no incidental identifying information is present, and have replaced user-names with pseudonyms.

Our data is drawn from the fifty-five forums that were labeled as “featured” on WebMD at time of data-collection. Although WebMD hosts other forums, the featured forums were administered by WebMD staff, and were monitored for spam and other inappropriate activity. The dataset spans a period from January 2009 to August 2014, and includes roughly one million posts and roughly 275 k unique users. However, as we describe below, our algorithmic approach to identifying MCs only yielded results for 23 forums, reducing the total size of the analyzed data set to 225,577 posts and 73,379 users.

2.2. Methods

As discussed, our primary aim was to identify and examine longitudinal sequences of RoA. We pursued this aim with two research questions:

RQ1. Can we identify longitudinal sequences of RoA across discussion threads in an OHC, and if so

RQ2. What kinds of RoA can we identify?

To address the first research question, we developed a computational algorithm that extracts MCs using the operationalization

provided above. Lacking a clear theoretical basis for parameters k and s we applied established data-analytic methods for estimating them (Goggins, Mascaro, & Valetto, 2013; Goggins, Valetto, Mascaro, & Blincoe, 2013). We then used crowd-based content analysis to determine how well people could distinguish between MCs and non-MCs given an intuitive description of RoA. To address the second question we applied a qualitative content analytic procedure to identify patterns in MCs that reflect RoA. The following subsections describe each of these methods in more detail.

2.2.1. An algorithmic approach to extracting MCs

Our approach to extracting MCs involved several steps:

- 1) We developed a machine-based classifier to identify those posts that provided or requested information support.
- 2) We used a dynamic topic analysis procedure to label the topics in these posts.
- 3) We implemented the algorithm described above, using the topic labels assigned in (2).
- 4) We set parameters s and k for each individual forum by looking for natural breaks in the data.

Each of these steps is described more fully below.

Labeling Information Support: Given existing concerns about the quality of information in online advice, we sought to restrict our analysis to posts that contained *informational* support (Bambina, 2007; Wang, Kraut, & Levine, 2012)—i.e., advice containing some sort of clinical guidance, rather than purely emotional support or simple banter. To identify informational support we randomly sampled 200 posts from each forum and, following Introne et al.'s (2016) approach, hired Amazon's Mechanical Turk to classify posts as providing or requesting informational support. We required that all coders be based in the US, to have completed at least 100 HITs, and to have at least a 98% approval rate. Five coders labeled each post, and they were paid at a rate intended to achieve at least minimum wage in the requester's home state.

We used intra-class correlation (ICC) to evaluate the consistency of the coders, which is appropriate when coders may vary from item to item (Koch, 2004; Wang et al., 2012). We obtained values of 0.90 for the “request information” category, and 0.93 for the “provide information” category, indicating an excellent level of agreement in both cases.

We used this set of posts to develop a classifier for extending the labeling to the entire dataset. We evaluated a number of classifiers, and obtained the best performance with a log regression-based classifier. We used four features for each post, as follows:

1. The depth of the post, with 0 indicating that the post was a top-level post.
2. A single binary feature, indicating whether or not the author of the post was the initiating poster in the thread.
3. The number of sentences terminating in some set of question marks.
4. A term-vector based representation of the post, using raw counts of each term. Stop-words were not excluded, but terms were stemmed and required to be at least three characters, and only non-numeric characters were included.

The classifier obtained mean accuracy scores of .90 for “request information” and 0.85 for “provide information” under 10-fold cross-validation, which are generally considered to be very high scores (García Adeva, Pikatza Atxa, Ubeda Carrillo, & Ansuategi Zengotitabengoa, 2014). We subsequently applied the classifier to the entire dataset.

The number of posts providing and requesting information were well correlated across the forums (mean $R = 0.88$). On average, information support posts (either providing or requesting information)

made up about 60% of all posts, but there were significant variations, with a minimum of 21% (the “relationships and coping” forum) and maximum of 84% (the “skin problems and treatment” forum). In total, roughly 524 k posts were labeled as containing informational support, and these were used in the subsequent topic modeling procedure.

2.2.2. Topic modeling

We applied topic modeling to all posts labeled as either providing or requesting informational support. Because we anticipated that advice would change over time, we sought to use a dynamic topic modeling procedure. Blei and Lafferty’s (2006) dynamic extension to the Latent Dirichlet Allocation (LDA) algorithm has been widely applied, but there are many parameters and some uncertainty about how best to determine the number of topics (Chang, Gerrish, Wang, Boyd-graber, & Blei, 2009).

Recent work (O’Callaghan, Greene, Carthy, & Cunningham, 2015) has shown that non-negative matrix factorization (NMF) (Lee & Seung, 1999) may produce topics with a higher degree of internal coherence than LDA when applied to corpora that include niche or non-mainstream content, such as might be expected in an online community that is driven by the emergent health needs of its users. Moreover, Greene and Cross (Greene & Cross, 2017) described a fully automated approach to generating dynamic topic models using NMF.

After applying this procedure to each of the 55 forums in our dataset, we discovered a total of 684 topics, with a large degree of variance in the number of topics among the individual forums: sixteen forums had between 18 and 20 topics, whereas the other forums covered a range from no topics up to 18 topics. In the three smallest forums (all ~1000 posts or less), no topics could be found, and these forums were omitted from further analysis. Among the remaining forums, 502 k of the informational support posts (96%) were labeled with a topic, and these were then used as input in the final steps of our analysis, described below.

2.2.3. Parameterizing time constraints for MCs: the cutoff value k

The parameter k reflects the intuition that after some period of time, a person is unlikely to carry information from one thread to another. Whereas others have analyzed the temporal dynamics of online discussions (e.g. Lampe & Resnick, 2004) and information diffusion processes (e.g. Simmons, Adamic, & Adar, 2011), there is little psychological work to guide the value of k in this context. Thus, we followed a data-driven approach, inspired by approaches to estimating web-session length (Arlitt, 2000) and other studies of online discussions (Goggins, Mascaro, et al., 2013; Goggins, Valetto, et al., 2013), and extrapolated a value for k using inter-post intervals in thread-based conversations on a per forum basis.

For each forum, we examined the distribution of time elapsed between subsequent posts in a single thread for any given individual. This interval is relevant because it allows us to establish a maximum beyond which it is unlikely a person will return to a typical conversation. We found that different forums exhibited different ranges of values, but distributions in all cases were distinctly bimodal, with one mode peaking at well less than an hour (~10 min) and the other at roughly one day (see Fig. 2). Examining the data more closely, it became apparent that individuals would respond to multiple prior posters or quickly follow-up on their previous responses, in “sessions.” We screened out these “single session” posts, by eliminating all but the first of any sequence of posts by a single user (with no intervening posts) that were separated by less than an hour.

After transforming the data in this way, we found inter-post intervals to be log-normally distributed in each forum. Thereafter, we set k for each forum to be two-standard deviations above the mean in the log-transformed data. k varied from roughly two weeks to nearly 300 days across the forums, but was itself log-normally distributed with a mean of about 57 days.

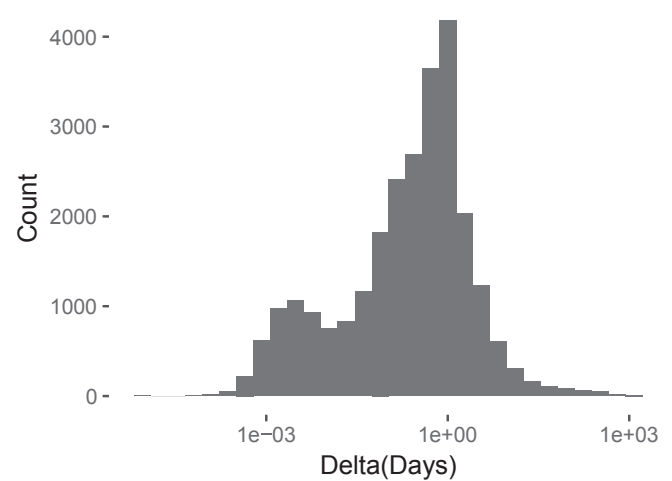


Fig. 2. Distribution of inter-post intervals for users in threaded discussions for the diabetes forum.

2.2.4. Parameterizing similarity for identifying MCs: Determining the cutoff value s

We developed a simple script (available at [URL omitted for blind review]) to extract MCs (as graphs of posts) using the previously developed topic assignments and values for k . We labeled each edge of the extracted graphs with the Jaccard similarity of the source and target posts (Gomaa & Fahmy, 2013). Specifically, Jaccard similarity here is the ratio of the size of the intersection to the size of the union of unique words in each post, after stop-words have been removed and words have been stemmed.

We then analyzed these graphs at different levels of s in order to determine an appropriate cutoff in each forum. As the cutoff value s is varied from 0 (indicating that two posts on each edge need not share any terms) to 1 (indicating that two posts on an edge have exactly the same set of words), the induced graphs degrade into an increasing number of small, weakly connected components. A *weakly connected component* is a graph-theoretical term indicating that there is a path (ignoring direction) through a graph connecting any pair of nodes in that graph.

As an example, Fig. 3 visualizes the number of weakly connected

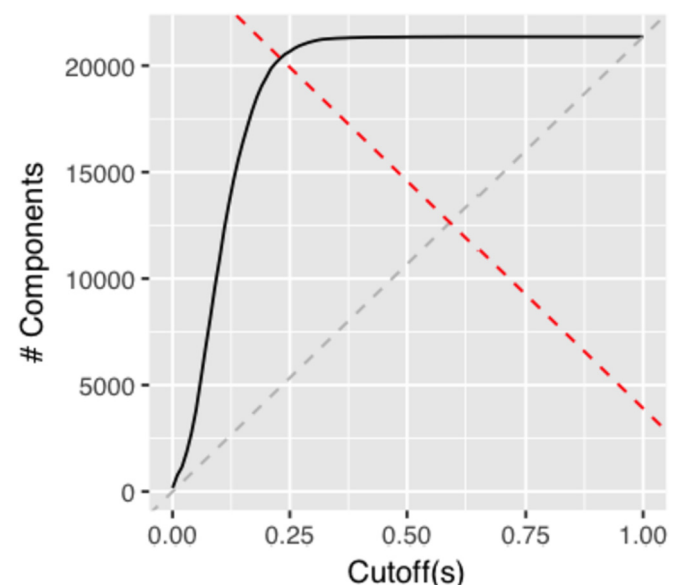


Fig. 3. Using the elbow method to determine a cutoff value for s in the diabetes forum.

components (where a connected component is presumed to be an MC) in the diabetes forum as s is varied from 0 to 1. The curve exhibits a distinctive “elbow” at which point the slope of the curve changes rapidly, and this becomes a natural point for the cutoff value of s (c.f. Solomon & Wash, 2014). To determine a precise value for s , we identify the point on the curve with the maximum perpendicular distance to a line between the minimum to the maximum values. In the diabetes forum, this value is 0.23, indicating that posts must share almost one quarter of their words to be considered to be connected. We found s to be normally distributed across the forums, ranging from 0.12 to 0.31, with a mean of about 0.19.

For our analysis, we considered each weakly connected component in a forum, setting parameters k and s as described above, to be an MC. After transforming the data in this way, we found that only 23 of the forums exhibited MCs with a maximum path length that was longer than two posts (recall, a single link is an MC lies between two posts, even though three posts are required to infer its existence). The remaining forums were smaller forums, with smaller (or non-existent) cores (as reported by Introne et al. (2016)). After some examination of our results, we eliminated these smaller forums from further consideration, because they did not appear to exhibit RoA. We revisit this choice in the discussion section of the paper.

2.2.5. Validating longitudinal reification of advice (RoA)

To determine if our MCs indeed seemed to carry sequences of RoA, we ran a manual content analysis task with sets of posts that were drawn from MCs and different groups of posts. Our rationale for this validation procedure was that if MCs indeed capture instances of RoA, a pool of non-experts should be able to distinguish them from other sets of posts not labeled as MC's on the basis of a relatively intuitive explanation of advice reification, as follows:

- Posts contain at least some advice that is roughly about the same kind of “thing” (though they may also contain other information).
- Posts are organized sequentially, and advice may change incrementally. That is, advice may be elaborated, or blended with other pieces of information, or become more refined.
- Posts are always taken from different conversation threads (though posters in meta-conversation posts may address one another).

To construct a set of MCs, we first drew a set of up to ten non-overlapping MCs from each of the 23 forums. Note that while the full MCs are in fact directed acyclic graphs, we felt that it would be cumbersome to explain and present these to our coders, and so selected paths from MCs, so each path was a strictly sequential set of posts. Based on our initial inspection of MCs, we excluded paths that were smaller than three posts, because we felt that two posts contained too little information to make a reliable determination about RoA. Finally, to construct our test set, we randomly selected paths through MCs in each forum in round robin fashion from the largest components down. For instance, if a forum had two components of length seven, twelve of length four, and twenty of length three, our final set of ten posts for that forum would consist of both paths of length seven, four of length four, and four of length three. We did this because smaller components were over-represented in the data (see below) and we felt that it was important to examine both longer and shorter MCs. We refer to this set of posts as the MC set.

To build a comparison set, we sought to construct sets of posts that might have some, but not all of the features of MCs. We did not want to compare MCs to random sets of posts because the lack of topical coherence would be an obvious distinguishing feature. Thus, for each forum, we constructed two additional sets of non-MC based paths in order to match the path lengths in the set of MCs. In the first set, we relaxed the assumption that Jaccard similarity had to meet the cutoff value s . More specifically, we modified Eq.(3) so that $Sim(P_{i1}, P_{i2}) < s$. We refer to this set as the *NotSim* set. In the second set, we altered the

requirement that posters had to appear in the same thread within time window k preceding a given post, and instead required that posters could not have appeared in the same thread within time window k . We refer to this set as the *NotComponent* set, reflecting the fact that these posts did not lie on the weakly connected components identified by our algorithm. Note that in both sets, we retained the restriction that posts in set all had to have the same topic label.

In total, this procedure produced 604 candidates to be labeled in our content analysis task. We ran this task using Amazon's MTurk service, with five MTurk workers per post. We paid workers in a manner consistent with minimum wage in the requester's home state. We only allowed workers with 98% or better approval rate who were based in the US, and we also asked workers to pass a qualifying task first, with ten different sets of posts that we hand-selected from our three sets. In our instructions for the qualifying task, we explained that some sets of posts were MCs, and that others were not, and provided workers with the definition above. We required that each poster get a perfect score on the qualifying test, and in the end invited ten people to code the complete set.

2.2.6. Identifying genres of reification of advice (RoA) with qualitative analysis

To address the question of what kind of RoA appears in MCs, we pursued a qualitative investigation of the extracted MCs to identify candidates for advice reification as it relates to the COP literature. Our overarching goals in this process were to provide some further confirmation that MCs are venues in which RoA occurs, while at the same time developing deeper insights about the manner in which advice is indeed reified.

Methodologically, our approach blends elements from grounded theory and conversation analysis, but does not fit precisely in either category. As with conversation analysis, our unit of analysis is a set of discourse moves, and we were attuned to regular structures and interaction patterns made visible in the text itself, rather than specific themes in the text. As with grounded theory, we (the authors) first worked independently to identify visible patterns in MCs that seemed to illustrate advice reification. We began with an orientation to RoA as it has been described by Wenger (Wenger, 1998), but also pursued the task as an iterative and inductive effort. After working independently, the authors worked together to develop a final set of categories, using exemplars identified in our individual examinations to help clarify these categories. We report on the categories we developed, and provide exemplars of each.

3. Results

3.1. RQ1: identifying sequences of RoA

We found that roughly 5% of all posts examined belonged to meta-conversations consisting of at least two posts, and component sizes were in general distributed as a long tail. A total of 6,578 meta-conversations were found. Here, we report several measures obtained from the crowd-based coding procedure to validate that these algorithmically identified sequences. First, we examined the Inter-Class Correlation (ICC) statistic, which as described above is appropriate for randomly assigned coders (Koch, 2004; Wang et al., 2012). Cicchetti (Cicchetti, 1994) suggests that ICC values below 0.4 are poor, from 0.4 to 0.59 as fair, 0.60-0.74 as good, and 0.75–1.00 as excellent. ICC analyses typically report values for both single raters, indicating the typical performance of a single rater, and an average value, reflecting the average performance of the group.

ICC2 for single raters was fair, at 0.50, and excellent for the overall group, at .85. This indicates that, given the definition provided above, raters were able to distinguish between the different MCs and non-MCs with a fair degree of reliability. This result provides us with a baseline validation that MCs are visibly different than other sets of posts.

We then examined the accuracy of the aggregate opinion of our coders to develop additional insights about the task. Overall, the *sensitivity* of the coders ($\#$ of true positives/ $\#$ of true positives + $\#$ of false negatives) was .82, and the *specificity* of the coders ($\#$ of true negatives/ $\#$ of true negatives + $\#$ of false positives) was 0.80, indicating that the aggregate measure erred in both directions by roughly the same amount. However, examining the percent accuracy across the three post sets is more revealing. For MCs, the aggregate response was correct 82% of the time. For *NotComponent* posts (those posts that were drawn from posters that did not post in the same threads within the window k), the aggregate response was correct 89% of the time, indicating that coders did not often identify features matching our definition of MCs in these posts. However, for *NotSim* posts, coders were correct only 67% of the time. As further confirmation of this result, we re-evaluated ICC omitting *NotSim* post-sets from the dataset; this raised the performance for single coders to 0.67 (indicating good reliability) and 0.91 for the group.

Our results indicate our informal definition of advice reification is sufficient to distinguish MCs from other sets of posts with a fair degree of reliability. This in turn provides good evidence that our algorithmic approach indeed extracts sequences of posts that reflect sequentially arranged instances of RoA. Our results also suggest that while the appearance of posters in the same threads within a fixed time window is a good indicator that people will perceive a set of posts to be a MC, the Jaccard similarity cutoff may be too conservative, and the degree of overlap in the linguistic content of posts may not be as important in establishing MCs. We return to this observation in our discussion section.

3.2. R2: genres of meta-conversation

In our qualitative analysis, we found a great deal of diversity in MCs. As a general observation, MCs often involved very long, highly informative posts, and were delivered by members who spoke with authority and confidence. In some MCs, there seemed little that appeared as RoA. However, we did encounter several distinctive genres¹ that were well represented in the data. In the following, we describe three of the clearest patterns. We provide several examples to support our observations; in all cases, we have replaced poster usernames with pseudonyms.

3.3. Boilerplate advice

With the boilerplate approach, some users would repeat large portions of advice to new users, tweaking it as necessary to make it relevant to context, and occasionally inserting new content that was apparently drawn from previous conversations. We observed this behavior across several different forums, and in some cases, different members in the same forum would cultivate different boilerplate (see Table 1).

The example in Table 3 is drawn from the Fibromyalgia forum. The content appears in many posts, and is often provided as a list of useful advice to welcome new members. The original author of the content is a member named YaYa. In the first post of the component, member Judy indicates that YaYa is not available, and re-posts YaYa's advice (which Judy had previously saved as a favorite). In the second post, Judy has returned to the forum, and posts a somewhat updated version of the content. Note, in YaYa's post (the second post in the component), the word 'despite' (highlighted in the excerpt) is now incorrectly spelled,

¹ Reciprocal and dynamic groupings are taken up in other fields as *genres* that are constituted of the situation, culture and universe of other genres (Devitt, 2004). Genre analysis recognizes that classification's fatal flaw is its inability to represent the different purposes and intents involved in the creation of a classification system, and a failure to recognize that the phenomena being classified is undergoing steady change.

and the new post contains both old and new content. Finally, in the last post, Judy re-posts just a verbatim portion of YaYa's updated post, and the spelling error is intact.

The spelling error highlights something about the plasticity of advice in this particular example. We queried our database for the misspelling 'dispite' and found that the vast majority of posts with that particular misspelling were in YaYa's content, and its earliest appearance was in a version of this boilerplate content that appeared at least a year before its appearance in this MC. Thus, Judy's initial correction of the word 'despite' indicates that she is not (at this point) merely copying YaYa's post, but is also exerting some editorial license. However, this change is minimal in comparison to those introduced by YaYa in the second post. We did not observe instances of these more significant alterations being made by other posters (as we might expect with a wiki).

The last post illustrates that Judy is clearly aware that updated content exists, and copies just a portion of it to address a more specific request. In the case of the thread where the post is found, the initial poster was not new to the forum, and made a specific request for YaYa's list. This is notable, because Judy had indicated that she had saved YaYa's earlier post as a favorite, indicating that either she updated the favorite or used an alternative method to keep track of YaYa's advice.

3.3.1. The memeification of advice

In all of the forums with pronounced MCs, we found evidence of advice becoming bundled into memelike phrases that were delivered across posts by multiple users. Such advice was not necessarily cut-and-paste (as with boilerplate advice) but there were distinct phrases suggesting that the community was developing a standard lexicon for presenting certain kinds of information. These phrases also became more compressed, echoing Simmons et al.'s (2011) observations about the compression of memes that spread on the internet. This compression might also reflect a process similar to that underlying the compression of referring expressions as collaborators develop common ground (H. H. Clark & Wilkes-Gibbs, 1986), and is a phenomenon Wenger invokes in his discussion of reification (Wenger, 1998, pp. 61–62;66).

One example of this occurred in the back-pain forum (Table 2). Two members, Beth and Bob, triaged the provision of advice across many different discussions in the forum, and both appeared in many of the MCs we examined. The selections illustrate the general evolution of advice about physiatrists both within and across MCs. Beth frequently recommended using seeing a physiatrist for managing pain issues. In component 2 (one of the earlier components identified in the forum), Beth provides advice about seeing a physiatrist, along with a link (which was still active at the time this manuscript was written) and a personal anecdote. Bob talks about pain, but does not explicitly mention physiatry. In component 3 however, Bob mentions physiatry in a context that is very similar to component 2. Beth also introduces the phrase "they go deeper into pain management." Note that Beth continues to provide the link to a web resource at this point. In component 4, which takes place roughly a year later, Bob has now adopted the phrase "they go deeper," and also includes the link that Beth had provided. Beth, however, has dropped the link in her advice; given that the link is still active, this is likely an editorial decision. By component 5, both Beth and Bob have dropped the link, and now use a fairly compressed phrase with the words "they go deeper."

3.3.2. Negotiated advice

A third type of MC reflected a more explicit discussion about specific information, transacted across multiple discussion threads by several people. These were among the most direct interactions we saw among MCs, even appearing to exhibit occasional adjacency pairs that are found in more direct conversational interaction (Sacks et al., 1974). In the excerpted MC (Table 3), posters in the diabetes forum discuss the proper interpretation and use of a sliding scale for an insulin regimen.

Table 1

Excerpts from boilerplate advice encountered in the Fibromyalgia forum. We have highlighted the word “despite” to draw attention to the change in spelling.

Component 1	
Judy, 9/13/2009 (542 words)	Hello and welcome to the FMily, I'm going to post a list from one of our beloved FMily members who is off the board for awhile while ill.: We are delighted to have you join us! You have certainly come to the right place for love, support, and understanding! You will find the greatest & most wonderful people here. You will get so much support, not to mention tips and all the great friends you will meet here! Its a wonderful place to be! We all have so much in common. You'll fit right in, my friend! You join in and share at any time you like. You will find someone here at all hours of the night as well! # The most important and most difficult is ACCEPTANCE! Then learning to live within your new found limitations. # Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength as a woman! If you lose that, then you lose so much of who you are, & what you stand for. I firmly believe in remembering all that you DO have to offer to others & the world, despite your current limitations! # Pacing yourself is a must, as well as getting restorative/restful nights sleep. # Taking hot baths helps to alleviate the pain and stiffness, many use heating pads and electric blankets... Hugs, Judy
YaYa, 9/27/09 (456 words)	Welcome to Our FMily! My name is YaYa, I'm an FM suffer of 18 years, along several autoimmune diseases. We are delighted to have you join us! You have certainly come to the right place for support! You'll learn tips & meet great people! Join in anytime! *The most important & most difficult is ACCEPTANCE! Then learning to live within your new found limitations. Setting realistic goals & making changes. Learning to say NO. *Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength! You are still YOU, just differently abled. I firmly believe in remembering/accepting all that you DO have to offer, despite your current limitations! *Pacing yourself is a MUST, even on good days. LISTEN to your body! Getting restorative/restful nights sleep is important. Rest & take breaks when your tired *Recommend U trying STOPAIN, it's very cooling/tingliny, to relieve pain. Found @ WalMart, in the pharmacy section...
Judy, 9/29/09 (407 words)	# The most important & most difficult is ACCEPTANCE! Then learning to live within your new found limitations. Setting realistic goals & making changes. Learning to say NO. # Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength! You are still YOU, just differently abled. I firmly believe in remembering/accepting all that you DO have to offer, despite your current limitations! # Pacing yourself is a MUST, even on good days. LISTEN to your body! Getting restorative/restful nights sleep is important. Rest & take breaks when your tired. # Recommend U trying STOPAIN, it's very cooling/tingliny, to relieve pain. Found @ WalMart, in the pharmacy section...

Table 2

Evolution of advice about physiatry in the back pain forum. Excerpts are from adjacent posts in selected components. Complete posts may be found in Appendix A.

Component 2	
Beth, 4/10/12	...may I suggest that you seek on that is a PHYSIATRIST Pain Management clinic - here's link on what they do etc., http://www.spine-health.com/treatment/spine-specialists/what-a-physiatrist I have been seeing one for about 4 years now - they have seen my records from before surgery, surgery and afterwards and what treatments i've had (which all have failed) - they are really good in my care...
Bob, 5/16/12	... The spine specialist may very well recommend your seeing a pain management specialist as well. Yes, there are definitely alternatives to pain medications and surgery. However, pain medications are sometime needed to reduce pain and increase functionality... You can read my story to see what pain management methods I am currently using successfully...
Component 3	
Bob, 12/27/12	... You will then probably be referred to a pain management specialist. The best pm doctors are physiatrists, as these doctors provide various pain management treatment for spinal problems and do no solely rely on prescription medication...
Beth, 4/9/13	...You can also see a PHYSIATRIST Pain Management specialist http://www.spineuniverse.com/treatments/what-physiatrist They go deeper into pain management control based on patients needs...
Component 4	
Bob, 4/13/14	...For pain management, a specialist that is a physiatrist is best as they go deeper and offer a wide variety of treatments. Here is a link to an article that explains what a physiatrist is and does: http://www.spineuniverse.com/treatments/what-physiatrist ...
Beth, 4/10/14	...Seeing a good pain specialist too would be beneficial A PHYSIATRIST Pain Specialist would be my preference as they go deeper into pain management by treating the underlying problems of pains and if previous treatments didn't work they won't try to repeat treatments...
Component 5	
Beth, 9/15/14	... You indicated you are going to a Pain Center for treatments. What type of doc's are they? Reason I am asking is a good pain doc specialist is a PHYSIATRIST as they go deeper into pain management control than regular pain centers. Please keep us posted what you find out...
Bob, 10/4/14	Also, if you see a pain management specialist the best are physiatrists. These go deeper and offer a wide variety of treatments. Please keep us updated.

To be clear, in this example all posts are taken from different threads, but lie on the same MC. In the first post, Amber provides some advice to Wendy, who is having difficulty with her insulin regimen. Amber describes the use of a sliding scale, which is a type of insulin therapy. There are various protocols for sliding scales, and each is tuned to a specific individual. Amber explains her regimen to Wendy, mentioning that she takes insulin with each meal based on her current blood sugar and carbohydrates consumed. Steve follows up, in a different conversation several hours later, and takes issue with how Amber is presenting the sliding scale. He addresses Amber directly, and argues that insulin should be taken as a correction following a meal. On the following day, John presents additional information that offers a potential resolution to the argument, pointing out that there are many differences in how sliding scales are used and discussed, highlighting the fact that any regimen is going to be specific to the doctor and patient. Finally, roughly a week later, Amber (speaking again to Wendy) re-emphasizes that the sliding scale is something that only a doctor can prescribe; however, she also includes the actual plan provided by her doctor, and uses it to illustrate all of the factors that come into play when determining actual insulin dosage.

3.4. Summary of meta-conversation genres

Our analysis reveals three different types of RoA. Each of these illustrate different points in negotiation about the advice presented. At one end of the spectrum, boilerplate advice is relatively static, de-contextualized, and uncontroversial; in Huh and Ackerman's words (2012), it is advice that has become 'polished' and stands on its own. At the same time, boilerplate advice is no longer actively shaped through participation—updates were not obviously the result of the collective efforts of forum members.

Memeification reveals advice that is more plastic, possibly becoming more static but still a target of negotiation. It is notable that this kind of advice is much shorter than boilerplate advice, but the phrase has a history and hence richer meaning for members than for newcomers. The potential for divergent interpretations of memeified advice across these different subpopulations highlights its position as a boundary object (Star & Ruhleder, 1996). The function of reification in creating such boundaries is also discussed at length by Wenger ((1998, p. 104). We observed many instances of phrases that appeared as stable, forum specific linguistic conventions that may have been the result of

Table 3
Excerpted posts from the diabetes forum, illustrating advice contention.

Component 5	
Amber, 2/14/09 1:53 p.m.	Hey Wendy, Everyone is trying to help you get a better understanding of your insulin meds. Just keep asking and keep talking. Not necessary to cry. We want you to stay healthy, that's all. So now you have insulin and an oral drug to help you with your blood sugar. Ask them exactly what sliding scale you are to stick to in writing. My doctor wrote out for me and explained while I was there having my A1c checked. Everyone is different, so this scale is written for YOUR needs, not anyone else's... The sliding scale is based on what your BS is before you eat and what you plan to eat at your meal. This depends on what your insulin ratio is to your carb ratio determined by your doctor. Ask for your sliding scale to be written down. I still have to look at mine daily to determine how much insulin I should take with each meal...
Steve, 2/14/09 5:33 p.m.	Amber, I have to correct you on the sliding scale. A sliding scale is not carb based. It is to lower an elevated bg AFTER meals.. The bg is taken aprox 2–3 h after a meal and then the correction is taken IF NEEDED. A sliding scale is only used for correction. If some are well controlled with the oral meds then most of the time the slideing scale does not even have to be used. You are using your slideing scale like a rappid acting insulin. This just does not work in a fast acting insulin. Your explanation is for a rappid acting insulin such as Novolog or Humalog. Insulins such as Novolin, or Humilin are the fast acting that is used on a slideing scale.
John, 2/15/09 1:36 p.m.	Carb Ratio and sliding scale are often used to mean the same thing, sometimes not, lack of STANDARD. In a way a carb ratio is a sliding scle, based on how many carbs in the meal at hand. The link you posted gave an example of a sliding scale based on BS levels, not food intake. My first sliding scale was based on urine sugars, so many units of NPH plus so many units of REGULAR for each plus level on the test-tape. With modern insulins newer diet CARB understandings etc. things are much better... A person on intensive therapy might be on a routine that involves testing before a meal AND taking any CORRECTION dose before a meal in addition to taking a after-meal insulin dose to cover the meal Routines for BS control do VARY, just as diabetics vary, not to mention that doctors themselves vary...
Amber, 2/23/09	Wendy, Your doctor should have given you your sliding scale. It's individualized depending on what your insulin/carb ratio is. Your blood sugar before you eat and what you are going to eat determine the sliding scale. Mine is 1:7. It used to be 1:15, but my sugar was way out of control at the time. It may change again. Here goes. It's what my doctor and CDE wrote out for me. As I said, this is individualized, so ask your doctor before you do anything. You may be giving too much insulin or not enough. Correction Factor Insulin for Carbohydrate Consumed BG Insulin Units Grms of carbs Insulin Units Less than 90 -2 7 1 90–120 0 14 2121–160 + 1 21 3161–200 + 2 28 4201–240 + 3 35 5241–280 + 4 42 6281–320 + 5 49 7 Slipper, do you see where I am going here? Before you bolus insulin for meals, you should ask yourself... what do I plan to eat? "What is my blood glucose now? "How active have I been before I eat? "How active do I plan to be after eating? In the carbohydrate consumed column, your ratio might be 1:15 or some other ratio. As you can see, the column on the left is in addition to what you are injectiing for the actual amount of carbs consumed.

such a process, and these may have important implications for advice as a resource for newcomers.

Finally, negotiated advice remains fluid and may be in the early stages of becoming reified by forum members. Negotiated advice bears some similarity to thread-based debates reported by Mamkykina et al. (2015), but here the debate proceeds less directly, as advice is delivered to others across multiple threads. Understanding whether and how individuals perceive the affordance of contentious conversations that are spread across multiple discussion threads is a valuable question for future study.

Although the different types of RoA discussed here can be organized along a continuum, we have no reason to believe they give rise to one another. However, they illustrate that different types of advice hold different status with respect to its stability within the community, and this is invisible to newcomers who might encounter it. In our closing discussion, we reflect further on this and what it might imply for future efforts to develop technical support for helping newcomers navigate OHCs.

4. Discussion & future work

With this paper, we have introduced an algorithmic method for extracting posts that are arrayed across multiple threads and demonstrated that these posts trace a continual process of advice reification. Our perspective is orthogonal to that taken in prior studies that examine collective sensemaking in OHCs, in two senses. It is literally orthogonal because meta-conversations move across the boundaries of designed affordances in the system (i.e. thread-based discussions). It is metaphorically orthogonal because it is theoretically crosscutting, highlighting the manner in which practice might evolve as a less intentional consequence of longitudinal member interactions.

Our analysis points to several deeper questions about OHCs and the forms that reification might take. Our algorithmic approach also paves the way for and examination of the relationship between the different types of advice and its quality, as well as design innovations to reveal some of the hidden context around advice that develops online. We consider these contributions, theoretical and applied, respectively in the following sections.

4.1. Trajectories of reified advice and collective intelligence

Wenger (1998) writes that the very act of articulation can be considered a type of reification, and this is perhaps even more so in online discussion forums where what has been posted persists indefinitely. Posting advice is particularly salient in an OHC because it reifies aspects of the central practice binding community members—namely, the management of a health condition. As we have shown, the same advice may be transformed in the interplay of participation and reification across a sequence of such posts, potentially varying with respect to its context dependence, degree of compression, and stability.

For Wenger, this interplay is a negotiation of meaning that transcribes the evolution of practice. In our case, the sequence of posts associated with each piece of advice localizes the evolution of a particular piece of practice within the history of the community. Borrowing the term ‘trajectory,’ which Wenger uses to describe transformations in member identity, we might consider trajectory of advice and its transformations directly as an object of study. Here, we have introduced a concrete approach to identifying trajectories in the reification of advice in OHCs. A fruitful course for future inquiry is to apply this as a theoretical lens, and examine an OHC as a bundle of such trajectories, woven through the shifting sociotechnical organization of the underlying community. This perspective leads us to frame our future exploration of OHCs around a nascent, mid-level theory of “Communities of Advice” (COAs). The goal of theorizing would be to explain how a dynamic sociotechnical system that is centered on people providing one another different forms of support (e.g., information, emotional) gardens advice without intentionality or explicit structural coordination mechanisms.

Evolving trajectories of advice create historical memory for an OHC, and as such, reflect a process of learning. Although this historical memory is at least partially shared by members—in the reified advice present in historical posts and the shared memories of past episodes of participation—Wenger for the most part locates learning within individuals. However, we propose that it is also productive to consider learning at the level of the overall sociotechnical system. From this perspective, an OHC can be regarded as a type of emergent collective intelligence which, over time, learns how to give advice to people that seek support for a given health condition.

Framing an OHC as a collectively intelligent system for providing healthcare advice orients us to a design-oriented research program to enhance learning along advice trajectories and make that advice more accessible. There are several incremental research questions that should be answered along this path. For instance, expanding on the work here, can we develop a more complete inventory of the dynamics of participation and reification along different advice trajectories across different OHCs. How might we quantify learning along these trajectories? On the basis of such research, might we build more effective member support for recalling and adapting previously reified advice? Can we help build features to support more resilient advice-giving practices in the face of continual member turn-over? For newcomers, can we offer signifiers that help them understand how community history may (or may not) anchor advice appearing in the posts they encounter?

Ultimately, these questions must run into those that have been raised about misinformation and advice in OHCs (e.g., Eysenbach et al., 2002), and we consider these more directly below.

4.2. Advice trajectories and quality

OHCs may provide advice seekers with good advice, but this can be intermingled with less valuable, or even misleading information (Eysenbach et al., 2002; Hartzler & Huh, 2016; Kanthawala et al., 2016). Core members may have contextual knowledge about forum participants and processes that enables them to distinguish between good and bad advice, and often correct misinformation in subsequent posts (Culver, Gerr, & Frumkin, 1997; Esquivel et al., 2006). However, outsiders and infrequent users may lack the context to recognize bad advice, and this could potentially have dire consequences (Venkatesan et al., 2013).

What to do about bad advice remains an open question. Huh and Pratt (2014) suggested adding automation to insert clinical expertise into forums, but such technology remains aspirational, and could be disruptive to the evolution of practice within a community. We believe that another fruitful line of inquiry would be to identify advice trajectories that exist and explore the use of these trajectories as proxy indicators of advice quality. Algorithms that leverage advice trajectories may also be easier to explain to end-users; such algorithmic transparency is an increasingly important sociotechnical concern (Shin & Park, 2019).

The algorithms we have developed to support the analysis presented herein offer the beginnings of such a solution by identifying connected instances of advice reification across threads. Our results are promising, but there remains much room for improvement. Our algorithm produces many MCs with very small path lengths, and we found little to suggest that these MCs contained useful examples of RoA. It is possible that our Jaccard similarity cutoff was overly restrictive or too simplistic a criterion, leading us to truncate paths prematurely. This would be consistent with our finding that coders were more likely to identify sequences of posts as MCs when they were on a component that did not meet the similarity threshold. However, our qualitative analysis also revealed several examples of linguistic convergence. There has been a fair amount of research developing online text-analytic procedures for measuring linguistic variation and convergence in online communities (Danescu-Niculescu-Mizil, West, Jurafsky, Leskovec, & Potts, 2013; Eisenstein, O'Connor, Smith, & King, 2010; Goel et al., 2016), and an analysis of emerging linguistic communities might serve as a more powerful approach than simple measures of token similarity.

A key remaining question is whether or not quality improves along an advice trajectory. There are good arguments to believe that this may be the case. One of these is drawn from theoretical work on collective intelligence. As Page (Hong & Page, 2004; Page, 2008) has argued, diverse groups of non-experts can outperform more homogeneous groups of experts, because they bring to bear different perspectives on a problem. In the context of an OHC, a group of individuals bring diverse lived experiences, and carry clinical information from many individual

doctors. Deliberative engagement in the construction of advice could plausibly produce novel collective insights that escape individual expert medical practitioners.

On the other hand, as Wenger points out, the reification of practice can also “ossify activity around its inertness” (Wenger, 1998, p. 61). In other words, the process of reification may freeze advice, and through repetition a piece of advice may begin to appear to the community as a truth that no longer needs careful examination. This can be especially dangerous in the case of healthcare advice, which can evolve and must often be tuned to the specific circumstances of individual patients.

Thus, another important future research challenge is to investigate the quality of advice contained in MCs. Following Huh and Pratt (2014), we anticipate that it will be important to evaluate some of the more subjective aspects of advice; that is, is the advice balanced, focused and reassuring? Does it direct individuals to clinical care when appropriate? At the same time though, we are interested in whether the process of reification improves the underlying information that is conveyed. Although many clinically oriented studies have lamented the existence of misinformation in online forums, misinformation is context sensitive (Huh & Pratt, 2014), the medical community may lack consensus about specific medical issues, and some misunderstandings may be more dangerous than others. These are important methodological challenges in future efforts to evaluate the quality of medical information in online forums.

4.3. Limitations

As an exploratory study, ours is limited in several ways. Regarding the generality of our findings, we anticipate that the existence of advice trajectories depends heavily on the existence of a group of core members who help sustain the institutional memory of the community. Our data happens to be drawn from a set of online health communities that have robust cores, but it is not clear that all OHCs exhibit this kind of social structure (though it is clear that some do).

More subtly, while we have uncovered sequences of posts that we believe reflect trajectories of advice reification, this does not conclusively demonstrate that these trajectories exist. That is, we have identified posts that fit an algorithmic description we believe is consonant with a conception of advice trajectories, but it may be that given enough data, it will always be possible to find posts that match these criteria regardless of whether they are meaningful aspects of a historical learning process for community members. This is a limitation of our methods, which are at their core quantitative and computational in nature. Future work would benefit from a natively qualitative approach that considers members' perspectives. Understanding how individuals perceive the social fabric of an OHC, and if they recognize shared resources such as remembered and reified advice would substantially enrich conversations about the existence of advice trajectories.

Despite these limitations, we believe that our analysis has opened a new avenue for understanding OHCs as sociotechnical entities that do more than merely satisfy the support needs of individuals. Through the lens of our analysis, an OHC is an emergent, collectively intelligent system that refines and adapts advice for health-condition management. The question we arrive at is how best to design social platforms that support this process as effectively as possible.

5. Conclusion

Online health support is an important tool in the evolving landscape of digital health. Yet, while many have rushed to embrace the role of big data and personal informatics, the expertise of those who successfully manage their own health conditions and support others is often marginalized, and even disparaged. We believe this expertise to be an untapped resource, especially when it is pooled and enhanced through its online expression in the form of advice that is freely given. Skepticism about peer-based support has been driven by the concern

that communities of individuals lacking clinical training might promulgate dangerous misinformation, and it is certainly the case that misinformation may be found on numerous health support platforms. However, it is also certainly the case that misinformation does not define these communities *en masse*.

In this paper, we have drawn on the concept of reification in communities of practice in order to surface some of the collective community processes that may adapt and refine advice over time. We have demonstrated an algorithmic approach to identifying connected sequences of advice reification, which we denote as advice trajectories, and described several different classes of advice trajectory. Our approach and findings enrich the conversation about OHCs and their value in several ways. We show how, at least in some cases, advice given by a “random stranger” online is neither the sole product of that individual, nor is its source uncritically assessed hearsay that bounces around the echo-chamber of the web. Instead, the advice we have focused on is a community resource, collectively processed over multiple instances of reification by a group of committed and engaged individuals. Our results immediately suggest that different communities likely develop different internal processes that alter the way this shared

resource develops and is managed over time.

These findings highlight a specific design challenge—namely, how can we design technical support and community processes so that members of an OHC can easily maintain and manage this community resource, without interfering with the fluid development of advice trajectories? We believe this is a more tractable challenge than attacking the problems of misinformation and bad advice directly. Pursuit of this challenge also includes an awareness of the internal community processes that may be instrumental in sustaining OHCs that operate as communities of practice. How to improve advice quality remains a central aim, but doing so should not come at the expense of the organic processes that have allowed OHCs to proliferate and thrive on the web.

Finally, our approach also introduces a novel path to revealing the social context of OHCs. With further development and testing, the approach could be readily deployed to existing sites. This would provide newcomers with a novel set of signifiers, enabling them to explore how advice may be anchored in the history of these evolving sociotechnical platforms, shaped and adapted by the practice of many members over time.

Appendix A

Full text of excerpted posts in identified meta-conversations.

Author/Date	Post Contents
Component 1: Fibromyalgia	
Judy, 9/13/2009 (542 words)	Hello and welcome to the FMily, I'm going to post a list from one of our beloved FMily members who is off the board for awhile while ill.: We are delighted to have you join us! You have certainly come to the right place for love, support, and understanding! You will find the greatest & most wonderful people here. You will get so much support, not to mention tips and all the great friends you will meet here! Its a wonderful place to be! We all have so much in common. You'll fit right in, my friend! You join in and share at any time you like. You will find someone here at all hours of the night as well! # The most important and most difficult is ACCEPTANCE! Then learning to live within your new found limitations. # Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength as a woman! If you lose that, then you lose so much of who you are, & what you stand for. I firmly believe in remembering all that you DO have to offer to others & the world, despite your current limitations! # Pacing yourself is a must, as well as getting restorative/restful nights sleep. # Taking hot baths helps to alleviate the pain and stiffness, many use heating pads and electric blankets. # Doing gentle stretches before or after getting out of bed, and after a hot bath helps to loosen tight muscles, taking deep breaths as you do them. # Massage & hot stone therapy is great for getting relief immediatly, although, it doesn't make it all go away. It relieves the intensity of the pain. It feels good and hurts at the same time. I personally, get a full body massage 2–3 x week or as needed. # Warm water aerobics is a gentle way of getting in exercise into your weekly routine. This helps to keep you mobile. You cannot feel your weight in pounds in the water, as you don't feel your pain. This feels really good. Most dr's highly recommend this form of exercise for FM patients. # Avoiding as much stress as possible, is important as to prevent some pain. # Continue educating yourself and family. # Learning to ask for help, can be very beneficial, as this will relieve you of some tasks others can do. # Eating a healthy diet. # Doing gentle exercises, & stretches are good habits to keep muscle strength and tone. Exercise body balls are great for stretching! # Staying connected with other people to avoid isolation & depression. # Exercise your brain daily with jigsaw puzzles, crossword puzzles, or sudoku to help keep your mental faculties stronger & to help brain fog. Look up a new word or two in the dictionary... Challenge your mind doing mental exercises to keep your mind sharp. Look at it as having a positive effect on your brain as a good walk does on your body. # When experiencing cognitive problems do one thing at a time to avoid concentration difficulties. # Keep “to do” notes in a designated area as reminders. Using handheld recorders or palm pilots to record “things to do” can be helpful for some. Below are a list of sites you can go to for more info on FM. 2008 Websites for FM info: www.fibromyalgia-symptoms.org/index.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp www.nfra.net www.fmaware.org (FibroAware Magazine & FREE Newsletter) www.healthcentral.com/chronic-pain/fibromyalgia.html www.healthcentral.com/chronic-pain/fibromyalgia-38174-5.html www.mayoclinic.com/health/fibromyalgia/DS00079/UPDATEDAPP=0 Hugs, Judy
YaYa, 9/27/09 (456 words)	Welcome to Our FMily! My name is YaYa, I'm an FM suffer of 18 years, along several autoimmune diseases. We are delighted to have you join us! You have certainly come to the right place for support! You'll learn tips & meet great people! Join in anytime! "The most important & most difficult is ACCEPTANCE! Then learning to live within your new found limitations. Setting realistic goals & making changes. Learning to say NO. "Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength! You are still YOU, just differently abled. I firmly believe in remembering/accepting all that you DO have to offer, despite your current limitations! "Pacing yourself is a MUST, even on good days. LISTEN to your body! Getting restorative/restful nights sleep is important. Rest & take breaks when your tired. "Recommend U trying STOPAIN, it's very cooling/tinglly, to relieve pain. Found @ WalMart, in the pharmacy section. "THERMIPAQ, theraputic hot/cold packs & lavender rice/flax seed hot/cold packs help to relieve painful sore muscles. "Taking hot baths & soaking in epsom salt helps to alleviate the pain & stiffness! Heating pad, electric blankets, DeepHeating WelPatch R also helpful. "Doing gentle stretches before & after getting out of bed, & after a hot bath helps to loosen tight muscles, taking deep breaths as you do them. "Massage & hot stone therapy is a great alternative for temp. relief. "Walking & warm water therapy is a gentle way of getting in exercise into your weekly routine. This helps to keep you mobile. You cannot feel your weight in pounds in the water, as you don't feel your pain.. Dr's highly recommend this form of exercise! "Doing gentle exercises & stretches are good habits to gain muscle strength. Pilates & exercise body balls R great for stretching! "Simplify your life. Avoiding/preventing as much stress in all areas of our lives as much as possible is key, as this fuels added pain.. "Following the recommendations of your dr & taking prescriptions as prescribe, responsibly. "Educate yourself & family. Keep up with new research. Print out info for friends/family. "Learning to ask for help, can be very beneficial, as this will relieve you of some tasks others can do. "Have a good support system. Staying connected with other people to avoid isolation & depression. "Exercise your brain daily to help keep your mental faculties stronger & to help brain fog by doing cross words games.. Challenge your mind doing mental exercises to keep your mind sharp. "Know that FM is not life threatening! Accept your symptoms wax & wane, but flares are transient & ultimately improve. "Try to let go of your fears & the “what ifs” & accept that life is different for you! Believe & know that U can still find joy . Sites for FM: www.fibromyalgia-symptoms.org/index.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp www.nfra.net www.fmaware.org www.healthcentral.com/chronic-pain/fibromyalgia.html www.healthcentral.com/chronic-pain/fibromyalgia-38174-5.html www.mayoclinic.com/health/fibromyalgia/DS00079/UPDATEDAPP=0 www.webmd.com/fibromyalgia/guide/default.htm chronicfatigue.about.com/od/whatisfibromyalgia/a/understandfibro www.nlm.nih.gov/medlineplus/fibromyalgia.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp Simple Explanation of FM ~ www.chronicfatigue.about.com/od/whatisfibromyalgia/a/understandfibro.htm Letter to the Healthy World from the

Land of Chronic Pain & Fatigue: www.fms-help.com/healthy.htm Science of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_science Causes of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_causes Symptoms of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_sympt

Judy, 9/29/09 (407 words)

The most important & most difficult is ACCEPTANCE! Then learning to live within your new found limitations. Setting realistic goals & making changes. Learning to say NO. # Acceptance of the “New You” is so important in keeping your self confidence, security, & inner strength! You are still YOU, just differently abled. I firmly believe in remembering/accepting all that you DO have to offer, despite your current limitations! # Pacing yourself is a MUST, even on good days. LISTEN to your body! Getting restorative/restful nights sleep is important. Rest & take breaks when your tired. # Recommend U trying STOPAIN, it's very cooling/tinglily, to relieve pain. Found @ WalMart, in the pharmacy section. # THERMIPAQ, theraputic hot/cold packs & lavender rice/flax seed hot/cold packs help to relieve painful sore muscles. # Taking hot baths & soaking in epsom salt helps to alleviate the pain & stiffness! Heating pad, electric blankets, DeepHeating WelPatch R also helpful. # Doing gentle stretches before & after getting out of bed, & after a hot bath helps to loosen tight muscles, taking deep breaths as you do them. # Massage & hot stone therapy is a great alternative for temp. relief. # Walking & warm water therapy is a gentle way of getting in exercise into your weekly routine. This helps to keep you mobile. You cannot feel your weight in pounds in the water, as you don't feel your pain.. Dr's highly recommend this form of exercise! # Doing gentle exercises & stretches are good habits to gain muscle strength. Pilates & exercise body balls R great for stretching! # Simplify your life. Avoiding/preventing as much stress in all areas of our lives as much as possible is key, as this fuels added pain.. # Following the recommendations of your dr & taking prescriptions as prescribe, responsibly. # Educate yourself & family. Keep up with new research. Print out info for friends/family. # Learning to ask for help, can be very beneficial, as this will relieve you of some tasks others can do. # Have a good support system. Staying connected with other people to avoid isolation & depression. # Exercise your brain daily to help keep your mental faculties stronger & to help brain fog by doing cross words games.. Challenge your mind doing mental exercises to keep your mind sharp. # Know that FM is not life threatening! Accept your symptoms wax & wane, but flares are transient & ultimately improve. # Try to let go of your fears & the “what ifs” & accept that life is different for you! Believe & know that U can still find joy . Sites for FM: www.fibromyalgia-symptoms.org/index.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp www.nfra.net www.fmaware.org www.healthcentral.com/chronic-pain/fibromyalgia.html www.healthcentral.com/chronic-pain/fibromyalgia-38174-5.html www.mayoclinic.com/health/fibromyalgia/DS00079/UPDATEDAPP=0 www.webmd.com/fibromyalgia/guide/default.htm chronicfatigue.about.com/od/whatisfibromyalgia/a/understandfibro www.nlm.nih.gov/medlineplus/fibromyalgia.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp Simple Explanation of FM~ www.chronicfatigue.about.com/od/whatisfibromyalgia/a/understandfibro.htm Letter to the Healthy World from the Land of Chronic Pain & Fatigue: www.fms-help.com/healthy.htm Science of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_science Causes of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_causes Symptoms of FM: www.fmaware.org/site/PageServer?pagename=fibromyalgia_sympt # Sites for FM: www.fibromyalgia-symptoms.org/index.html www.niams.nih.gov/Health_Info/Fibromyalgia/default.asp www.nfra.net www.fmaware.org www.healthcentral.com/chronic-pain/fibromyalgia.html www.healthcentral.com/chronic-pain/fibromyalgia-38174-5.html

Component 2: Back Pain Exchange

Beth, 4/10/12

Hi Jason - welcome to the support group - I saw that your pain management physician dropped you - I hope that you have found another one - if not then may I suggest that you seek on that is a PHYSIATRIST Pain Management clinic - here's link on what they do etc., <http://www.spine-health.com/treatment/spine-specialists/what-a-physiatrist> I have been seeing one for about 4 years now - they have seen my records from before surgery, surgery and afterwards and what treatments I've had (which all have failed) - they are really good in my care and will not suggest any treatments that previously failed to be done. Also, with previous pain clinics I have been to I also had 2 of the nerve blocks done without success - however there are lots of members who have had this done and have had great results - each individual is different and reacts differently to treatments etc., I hope the site helps answers questions on Physiatrist Pain Management and that you have found one that does help you. Take care - Beth

Bob, 5/16/12

Hello. The best approach to take with the spine specialist is to explain your symptoms and how they are affecting your life. He will determine the best treatment methods for you to try. The spine specialist may very well recommend your seeing a pain management specialist as well. Yes, there are definitely alternatives to pain medications and surgery. However, pain medications are sometime needed to reduce pain and increase functionality. It is best to be open to what your doctors recommend. Also, if a certain specialist is not treating you as you would like, you can find one that will. You are in charge of your body and health care. Be proactive and do your research. I am very limited as far as what medications I can take for my high levels of chronic pain. You can read my story to see what pain management methods I am currently using successfully. I pray you find answers and pain relief soon. Click on my user name or avatar picture to read my story. Blessings, -Bob

Component 3: Back Pain Exchange

Bob, 12/27/12

Hello, Karl and welcome. I am sorry you are experiencing so much pain from your accidents. Just a quick background, which you can also read by clicking on my username or avatar picture, I have been managing moderate to severe chronic back pain for about 35 years. I have never had surgery nor been in an auto accident. My pain is from degenerative and genetic problems with my spine. I am an avid proponent of chiropractic as it works for me and my family. Only “straight” chirporactic, though, not the gadgets such as the DRX 9000 you had had done. I get a regular chiropractic adjustment about one per month and my neurosurgeon told me that will keep me out of a wheel chair for a few more years. Many people are totally against chiropractors and may voice their opinions on that. You are being very wise with getting opinions from multiple spinal surgeons. Make sure you are seeing doctors that specialize in the spine, not just a general orthopedist or neurologist. Have you seen a pain management specialist? That would be another good doctor to see. The best pm doctors are physiatrists as they offer all forms of procedures and treatments along with prescription medications. I pray you can get answers and relief soon. Please keep us updated. Blessings, -Bob

Beth, 4/9/13

Hi Wesley - welcome to the support group - make sure you are seeing the right kind of doctor - either a Orthopedic Spine Specialist or Neurosurgeon Spine Specialist - they may order current MRI/CT Scan so they can see what's going on. You can also see a PHYSIATRIST Pain Management specialist <http://www.spineuniverse.com/treatments/what-physiatrist> They go deeper into pain management control based on patients needs. they can also examine you, request MRI/CT Scans and can refer you to a spine specialist as they work with several that are really great spine specialist. Only the Spine Specialist and/or Pain Management Specialist can answer you questions once they have examined you. Please let us know what you find out ~ ~ ~ ~ Click on my name or picture and read my story ~ ~ Take care ~ ~ God Bless ~ ~ ~ ~ Joy ~ ~

Component 4: Back Pain Exchange

Bob, 4/13/14

Hello and welcome. I am sorry you are in so much pain and fully understand. The best type of doctor to see for any type of spinal problem is a spinal orthopedic surgeon or a spinal neurosurgeon. Surgery should be your last option, but these doctors are the best for diagnosing problems and recommending an effective treatment plan. You should never have any type of treatment until you have an accurate diagnosis from a spinal specialist as you could be causing yourself more pain and problems. For pain management, a specialist that is a physiatrist is best as they go deeper and offer a wide variety of treatments. Here is a link to an article that explains what a physiatrist is and does: <http://www.spineuniverse.com/treatments/what-physiatrist> There are many sites on the internet that have good information about all spinal issues such as [spineuniverse.com](http://www.spineuniverse.com) and [spine-health.com](http://www.spine-health.com). Everyone is unique and their problems and symptoms are unique to them, but I have listed the best steps towards getting the relief that you need. A general physician just is not experienced enough to diagnose complex spinal problems like yours. You need to see a spinal orthopedic surgeon and/or a spinal neurosurgeon, then a physiatrist. While your regular doctor may review the MRI and see no problems, a spine specialist could find the cause for your symptoms. I pray you can find answers and relief soon. Please click on my username or avatar picture to read my story. Blessings, -Bob

Beth, 4/10/14

Hi luluyang Can u describe more in detail what type of leg and foot cramps u are having? I'm wondering if the problem is more of and underlying problem meaning u have something more going on with ur spine whereas a bulging disc could be pressing on the nerve root causing the pains. I had problems of going from one pain clinic to another starting over with procedures . Pain clinic docs saw \$\$ and I didn't know any better not knowing what was going on with my spine. I truly believe u need to get a second opinion from a spine specialist either Orthopedic or Neurosurgeon Spine Specialist and having a MRI or CT Scan done possibly with and without dye injection as this will show more if what's going on. Seeing a good pain specialist too would be beneficial A PHYSIATRIST Pain Specialist would be my preference as they go deeper into pain management by treating the underlying problems of pains and if previous treatments didn't work they won't try to repeat treatments unless they were not done properly or in the right area. I hope this helps please keep up posted ~ ~ ~ ~ Click on my name or picture and read my story ~ ~ ~ ~ Beth ~ ~

- citation.cfm?id=1870658.1870782.
- Esquivel, A., Meric-Bernstam, F., & Bernstam, E. V. (2006). Accuracy and self correction of information received from an internet breast cancer list: Content analysis. *BMJ British Medical Journal*, 332(7547), 939–942.
- Eysenbach, G., Powell, J., Kuss, O., & Sa, E.-R. (2002). Empirical studies assessing the quality of health information for consumers on the world wide web: A systematic review. *Journal of the American Medical Association*, 287(20), 2691–2700. <https://doi.org/10.1001/jama.287.20.2691>.
- Fox, S., & Purcell, K. (2010). *Chronic disease and the internet*. Retrieved from http://www.pewinternet.org/files/old-media/Files/Reports/2010/PIP_Chronic_Disease_with_toplevel.pdf.
- García Adeva, J. J., Pikatza Atxa, J. M., Ubeda Carrillo, M., & Ansuategi Zengotitabengoa, E. (2014). Automatic text classification to support systematic reviews in medicine. *Expert Systems with Applications*, 41(4), 1498–1508. <https://doi.org/10.1016/j.eswa.2013.08.047>.
- Goel, R., Soni, S., Goyal, N., Paparrizos, J., Wallach, H., Diaz, F., et al. (2016). The social dynamics of language change in online networks. In E. Spiro, & Y.-Y. Ahn (Eds.). *Social informatics* (pp. 41–57). Springer International Publishing.
- Goggins, S. P., Mascaro, C., & Valetto, G. (2013). Group informatics: A methodological approach and ontology for sociotechnical group research. *Journal of the Association for Information Science and Technology*, 64(3), 516–539.
- Goggins, S. P., Valetto, G., Mascaro, C., & Blincoe, K. (2013). Creating a model of the dynamics of socio-technical groups. *User Modeling and User-Adapted Interaction*, 23(4), 345–379.
- Gomaa, W., & Fahmy, A. (2013). A survey of text similarity approaches. *International Journal of Computer Applications*, 68(13), 13–18. <https://doi.org/10.5120/11638-7118>.
- Greene, D., & Cross, J. P. (2017). Exploring the political agenda of the european parliament using a dynamic topic modeling approach. *Political Analysis*, 25(1), 77–94. <https://doi.org/10.1017/pan.2016.7>.
- Hartzler, A., & Huh, J. (2016). Level 3: Patient power on the web: The multifaceted role of personal health wisdom. *Health informatics. Consumer health informatics* (pp. 135–146). https://doi.org/10.1007/978-3-319-19590-2_6.
- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), 16385.
- Huh, J. (2015). Clinical questions in online health communities: The case of “see your doctor” threads. *Proceeding of the 2015 conference on computer supported cooperative work. Presented at the computer supported cooperative work, Vancouver, Canada*.
- Huh, J., & Ackerman, M. S. (2012). Collaborative help in chronic disease management: Supporting individualized problems. *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 853–862). <https://doi.org/10.1145/2145204.2145331>.
- Huh, J., & Pratt, W. (2014). Weaving clinical expertise in online health communities. *Proceedings of the SIGCHI conference on human factors in computing systems CHI conference, 2014* (pp. 1355–1364). <https://doi.org/10.1145/2556288.2557293>.
- Introne, J., Semaan, B., & Goggins, S. (2016). A sociotechnical mechanism for online support provision. *Proceedings of the SIGCHI conference on human factors in computing systems. Presented at the SIGCHI, san Jose, CA*<https://doi.org/10.1145/2858036.2858582>.
- Johnson, G. J., & Ambrose, P. J. (2006). Neo-tribes: The power and potential of online communities in health care. *Communications of the ACM*, 49(1), 107–113. <https://doi.org/10.1145/1107458.1107463>.
- Johnston, A. C., Worrell, J. L., Di Gangi, P. M., & Wasko, M. (2013). Online health communities: An assessment of the influence of participation on patient empowerment outcomes. *Information Technology & People*, 26(2), 213–235. <https://doi.org/10.1108/ITP-02-2013-0040>.
- Jones, A., & Preece, J. (2006). Online communities for teachers and lifelong learners: A framework for comparing similarities and identifying differences in communities of practice and communities of interest. *International Journal of Learning Technology*, 2(2–3), 112–137. <https://doi.org/10.1504/IJLT.2006.010615>.
- Kanthawala, S., Vermeesch, A., Given, B., & Huh, J. (2016). Answers to health questions: Internet search results versus online health community responses. *Journal of Medical Internet Research*, 18(4)<https://doi.org/10.2196/jmir.5369>.
- Kimmerle, J., Thiel, A., Gerbing, K.-K., Bientzle, M., Halatchliyski, I., & Cress, U. (2013). Knowledge construction in an outsider community: Extending the communities of practice concept. *Computers in Human Behavior*, 29(3), 1078–1090. <https://doi.org/10.1016/j.chb.2012.09.010>.
- Kittur, A., & Kraut, R. E. (2008). Harnessing the wisdom of crowds in wikipedia: Quality through coordination. *Proceedings of the ACM 2008 conference on computer supported cooperative work* (pp. 37–46).
- Koch, G. G. (2004). Intraclass correlation coefficient. *Encyclopedia of statistical sciences*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/0471667196.ess1275.pub2/abstract>.
- Lampe, C., & Resnick, P. (2004). Slash(Dot) and burn: Distributed moderation in a large online conversation space. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 543–550). <https://doi.org/10.1145/985692.985761>.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788–791. <https://doi.org/10.1038/44565>.
- Malone, T. W., Laubacher, R., & Dellarocas, C. (2009). *Harnessing crowds: Mapping the genome of collective intelligence*. MIT Sloan School: MIT Working Paper No. 4732–09.
- Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction*, 12(2), 201–232. <https://doi.org/10.1145/1067860.1067864>.
- Mamykina, L., Nakikj, D., & Elhadad, N. (2015). Collective sensemaking in online health forums. *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3217–3226). <https://doi.org/10.1145/2702123.2702566>.
- O’Callaghan, D., Greene, D., Carthy, J., & Cunningham, P. (2015). An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42(13), 5645–5657. <https://doi.org/10.1016/j.eswa.2015.02.055>.
- Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Retrieved from <http://books.google.com/books?hl=en&lr=&id=hJRu408q1xwC&oi=fnd&pg=PR13&dq=scott+page+the+difference&ots=HBNL4ugtGW&sig=epJLkbeW44rM2ikoScpYpEVg2v4>.
- Parnin, C., Treude, C., & Grammel, L. (2012). *Crowd documentation: Exploring the coverage and the dynamics of api discussions on stack overflow*. Georgia Institute of Technology.
- Quinn, A. J., & Bederson, B. B. (2011). Human computation: A survey and taxonomy of a growing field. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1403–1412). <https://doi.org/10.1145/1978942.1979148>.
- Ridings, C., & Wasko, M. M. (2010). Online discussion group sustainability: Investigating the interplay between structural dynamics and social dynamics over time. *Journal of the Association for Information Systems*, 11(2), Retrieved from <http://aisel.aisnet.org/jais/vol11/iss2/1>.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696. <https://doi.org/10.2307/412243>.
- Shin, D., & Park, Y. J. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 277–284. <https://doi.org/10.1016/j.chb.2019.04.019>.
- Silva, L., Goel, L., & Mousavidin, E. (2009). Exploring the dynamics of blog communities: The case of MetaFilter. *Information Systems Journal*, 19(1), 55–81. <https://doi.org/10.1111/j.1365-2575.2008.00304.x>.
- Simmons, M. P., Adamic, L. A., & Adar, E. (2011). Memes online: Extracted, subtracted, injected, and recollected. *Fifth International AAAI Conference on Weblogs and social media. Presented at the fifth International AAAI conference on weblogs and social media*.
- Solomon, J., & Wash, R. (2014). Critical mass of what? Exploring community growth in WikiProjects. *Proceedings of the eighth International AAAI conference on weblogs and social media (ICWSM)*. Presented at the ICWSM ’14. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8104>.
- Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research*, 7(1), 111–134. <https://doi.org/10.1287/isre.7.1.111>.
- Venkatesan, S., Han, W., Kisekka, V., Sharman, R., Kudumula, V., & Jaswal, H. (2013). Misinformation in online health communities. *Proceeding of the 2012 workshop on Information security and privacy. Presented at the pre-ICIS workshop on information security and privacy*. Retrieved from <http://aisel.aisnet.org/wisp2012/28>.
- Wang, Y.-C., Kraut, R., & Levine, J. M. (2012). To stay or leave?: The relationship of emotional and informational support to commitment in online health support groups. *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 833–842). <https://doi.org/10.1145/2145204.2145329>.
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.