



# Attraction, selection, and attrition in online health communities: Initial conversations and their association with subsequent activity levels

Babak Abedin<sup>a,b,\*</sup>, David Milne<sup>b</sup>, Eila Erfani<sup>b</sup>

<sup>a</sup> Macquarie University, Macquarie Business School, Sydney, Australia

<sup>b</sup> Faculty of Engineering & IT, University of Technology Sydney, Sydney, Australia

## ARTICLE INFO

### Keywords:

Online health communities  
OCASA theory  
Initial posts  
Activity levels

## ABSTRACT

**Background:** The effectiveness of online health communities (OHCs) for improving outcomes for health care consumers, health professionals, and health services has already been well investigated. However, research on determinants of OHC users' activity levels, what is associated with attrition or attraction to these communities, and the impacts of initial posts is limited.

**Objectives:** We sought to explore topic exchanges in OHCs and determine how users' initial posts and community reactions to them are associated with their subsequent activity levels. We also aimed to extend the theory of Attraction-Selection-Attrition for Online Communities (OCASA) to this area.

**Methods:** We examined exchanges in a major Australian OHC for cancer patients, analyzing about 2500 messages posted over 2009–18. We developed a novel annotation scheme to examine new members' initial posts and the community's reactions to them.

**Results:** The annotation scheme includes five themes: informational support provision, emotional support provision, requests for help, self-reflection & disclosures, and conversational cues. Initial conversations were associated with future activity levels in terms of active posting versus non-active engagement in the community. We found that most OHC members disclosed personal reflections to bond with the community, and many actively posted to the community solely to provide informational and emotional support to others.

**Conclusion:** Our work extends OCASA theory to bond-based contexts, presents a new annotation scheme for OHC support topics, and makes an important contribution to knowledge about the relationship between users' activity levels and their initial posts. The findings help managers and owners understand how members use OHCs and how to encourage active participation. They also suggest how to attract new members and minimize attrition among existing members.

## 1. Introduction

Interest in online health communities (OHCs) has grown substantially over recent years, particularly since 2007 [1,2]. Scholars have examined OHCs' potential to improve outcomes for health service users [3], health professionals [4], and health services and support [5]. An OHC is an information system and internet-based text forum in which people share health-related information, experiences and feelings, and provide support and encouragement to other community members [1,4,6]. Interactions in these communities can supplement traditional communications between doctors, patients and caregivers, and often the user-generated health content in these communities provide information and support that is not available elsewhere [7].

Previous studies have discussed users' motivations for joining OHCs,

such as ongoing management of health conditions [7], exchanging support and learning from other's experiences [3], accessing information and emotional help about rare and socially stigmatized health conditions [8], and forming close relationships with like-minded people (Fan & Lederman, 2019). Yet, there is still debate about how individuals use the community, what attrition means in bond-based communities like OHCs, and what can influence community members' active contributions to the community [9].

Researchers have already stressed the importance of initial posts. For instance, Burke et al. [10] found that new posts asking for help or informational support are a way of gaining membership of a group and entering a supportive conversation. Other scholars like McInnis et al. [11] and Levine et al. [12] also stressed on the importance of newcomers, and found that initial posts can have intentional or

\* Corresponding author at: Balaclava Rd, Macquarie Park, NSW, 2109, Australia.

E-mail addresses: [Babak.Abedin@mq.edu.au](mailto:Babak.Abedin@mq.edu.au) (B. Abedin), [david.milne@uts.edu.au](mailto:david.milne@uts.edu.au) (D. Milne), [eila.erfani@uts.edu.au](mailto:eila.erfani@uts.edu.au) (E. Erfani).

<https://doi.org/10.1016/j.ijmedinf.2020.104216>

Received 24 March 2020; Received in revised form 2 June 2020; Accepted 6 June 2020

1386-5056/ © 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

unintentional impacts to the dynamic of OHCs. In particular, Levine et al. [12] found motivations of newcomers are key in understanding the impact of initial posts on the community dynamic. Furthermore, Smithson et al. [13] have reported that often OHCs members initiate a thread by asking for ‘advice’ that had an obvious ‘safe’ answer for it. These scholars found that new posters often use self-disclosure as a way of entering membership of the community.

We draw on and contribute to the underutilized theory of Attraction-Selection-Attrition (ASA) for Online Communities (OCASA). The theory is comprised of three stages that explain user behaviors in online communities [14]: *attraction* (expectations and goals that encourage people to join the community), *selection* (how they use the community and whether the community serves their purpose), and *attrition* (people leaving the community because their needs were not satisfied). While the theory offers a unique perspective on user experience in online communities, it relies on a cost-benefit analogy for commercial purposes. Butler et al. [14] acknowledged that the theory needs to be extended to bond-based online communities (e.g. OHCs) in which members have altruistic and non-financial motivations to participate in the community. They invited future researchers to fill this gap, and stressed that “alternative operationalizations of the OCASA theory would strengthen its value as a foundational theory for describing the role of community platform technologies in community growth and sustainability” (ibid).

Based on OCASA and past research in OHCs (e.g. [1,12]), users’ activity level is a significant determinant of attractiveness, usefulness, and growth of any online community. While substantial research has already been conducted on various factors that may facilitate or influence users’ activity level in OHCs (e.g. [4,5,15]), little has been done to understand the impact of initial posts by new members and subsequent community reactions [11]. Lack of such understanding can lead to new members’ losing interest in the community or can damage community’s effectiveness [13,14]. Thus, in this paper we address this gap in the literature and hypothesize that the volume and topics of initial posts made by new members (H1), coupled with the community reactions to these posts (H2), can predict whether or not a new visitor will go on to become an active community member (Fig. 1). While we acknowledge factors such as members’ health conditions, living environments and other factors that may influence their activity level in OHCs, they often cannot be captured directly from textual discussions in the community, and thus excluded from the analysis in this paper.

In response to Butler et al.’s [14] invitation, the first objective of our research was to develop an annotation scheme for support topics exchanged between community users and to distinguish active and non-active posters in OHCs. The scheme allowed us to examine H1 and H2

and answer three questions: what topics are discussed in the initial conversations of new OHCs users? Does the volume of activity in a new visitor’s conversation influence their active or non-active status? Do the topics discussed in a new visitor’s initial conversation influence whether they become active or non-active? Our second objective was to extend the OCASA theory to a bond-based health context, in which members do not share a cost-benefit motivation for participation in the community. This is important, because despite the theory’s potential for understanding OHCs, it has not yet been widely exploited in altruistic contexts.

## 2. Research design

Content analysis of textual discussion is popular among OHCs scholars, as it offers rich insights about community members’ needs, interests and activity pattern [2,8,16]. As outlined in the following, we study an active OHC in Australia, and conducted a qualitative thematic analysis to extract and examine topics of discussions and to flag initial posts made by new members. We then conducted a hermeneutic approach for developing an annotation scheme, and use it to annotate conversations using a web application developed specifically for this study.

### 2.1. The study context

We studied an OHC owned and moderated by Cancer Council New South Wales (CCNSW). CCNSW is a member of Cancer Council Australia, and an independent charity, 94 % community funded. As Fig. 2 shows, it hosts a set of active online communities that allow people with cancer, and their carers, to exchange support with other community members (<https://www.cancercouncil.com.au/OC>). The community management and planning team was comprised of three experts who acted as external advisors of this research and provided feedback about the research direction and findings. Moderators rarely intervene in the conversations of community members to encourage peer-led conversations. Given this, moderators’ comments were rare and thus not examined in this study.

### 2.2. Data collection

We studied *I have cancer* forum (Fig. 2), a sub-community dedicated to cancer patients to extract themes of support exchanges. This ensured that the themes would represent a coherent set of content related to this group of OHC users. The data collection process began by first identifying all ordinary community members (i.e. those not formally affiliated with CCNSW) who made their first post on this sub-community between the 1 st of April 2009 and the 1 st of July 2018—859 users in all.

The users were classified as either *active* (members who regularly post in various discussion threads) or *non-active* (members who may only engage in one thread or occasional discussions). Given that we only had access to textual content, members’ activity level could only be captured based on the posts they made to discussion threads. In consultation with the CCNSW community management team, we considered *active* members as those who participated at least once in at least two conversations (i.e. on different threads) in the window starting 30 days (i.e. approximately one month) and ending 180 days (i.e. approximately six months) after the date on which they posted their first message on the forum. Otherwise, they were considered *non-active* members. This resulted in 159 *active* and 700 *non-active* members.

For each user, we gathered the conversation stemming from their first post: namely, their first message and all other messages posted within seven days in the same thread. These messages included messages written by the identified user (i.e. the *conversation starter*), by affiliated community members (i.e. *replying moderators*) and other un-affiliated community members (i.e. *replying peers*). The 859 users/

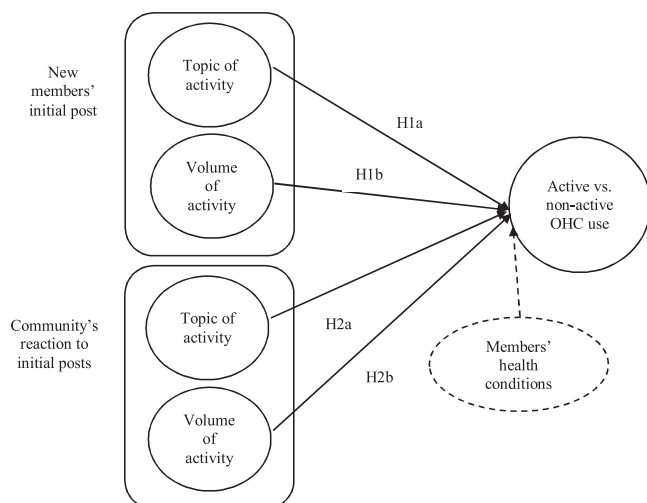


Fig. 1. Research Model.

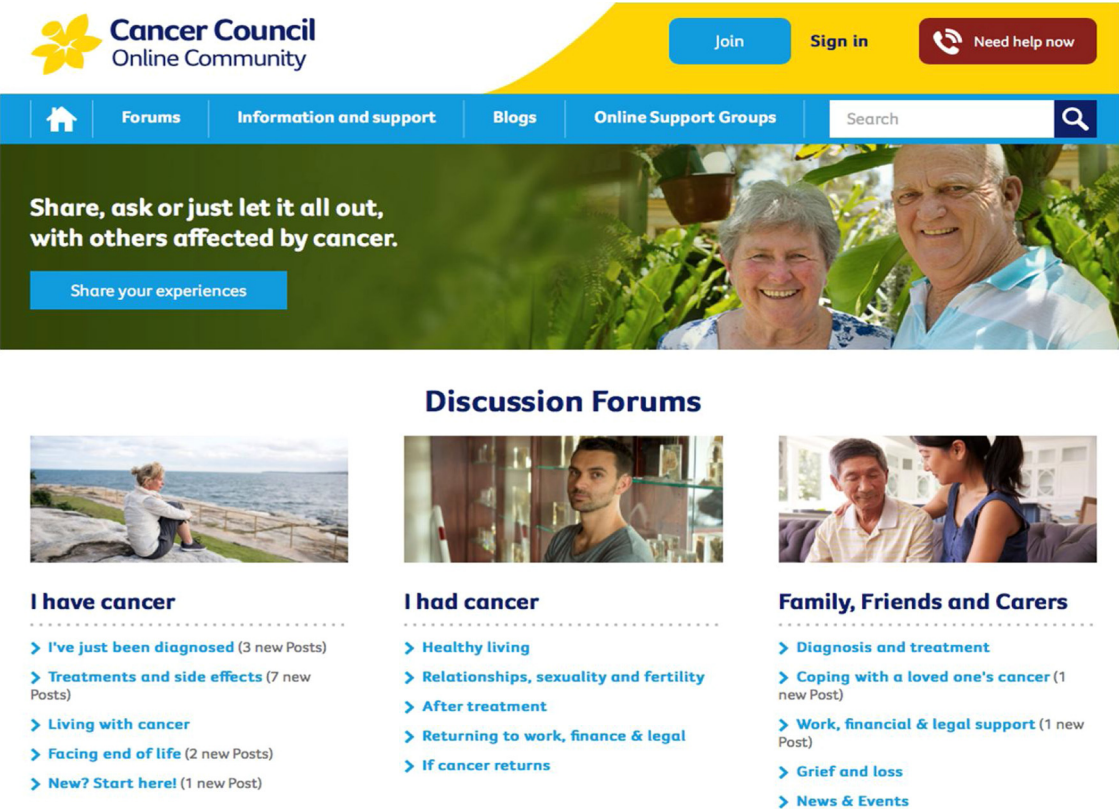


Fig. 2. Cancer Council NSW Online Communities.

conversations yielded a total pool of 2463 messages available for analysis.

We measured ‘volume of activity in a new visitor’s conversation’ by analysing the number of sentences as well as the number of messages written by the new members. We then measured ‘community reactions to initial posts’ by extracting the number peers replying to the new member’s posts as well as the number of sentences and the number of messages written by peers.

2.3. Development and validation of the annotation scheme

We employed a passive observational perspective in conducting a thematic analysis of the 2463 Cancer Council NSW community messages (Fig. 3). We used a hermeneutic approach for conducting a literature review [17] for developing the annotation scheme for practical research [17].

Step one, *Understanding* included searching the literature for an analytic development of key concepts (i.e topics discussed online), which led to five themes based on their similarities. We also used the methods outlined in Fahy [18] and Abedin et al. [19], which acknowledge topics related to social and non-task related discussions, to capture generic content and greetings such as salutations and rhetorical

questions (Table 1). Next, in step two, *Explanation*, we randomly selected small samples of messages from the forum and annotated them using the classifying scheme generated in the previous step. This ensured all sentences (and the corresponding meaning) and their corresponding meanings could be captured by the classifying scheme. Authors separately and manually annotated the text, and then met to discuss agreements and resolve disagreements. Secondly, we presented the annotation scheme to the three CCNSW experts to examine face validity of the instrument in order to assure integrity, meaningfulness, relevance, and clarity of the scheme. We conducted two meetings with CCNSW team, and incorporated their feedback about naming of themes as well as meaningfulness of the support topics and their relevancy to the CCNSW community. The last step, *Interpretation*, produced an annotation scheme of 15 discussion topics grouped in five themes (Table 2): Informational Support Provision (ISP), Emotional Support Provision (ESP), Request for Support (RS), Conversation Cues (CC), and Self-Reflection & Disclosure (SRD).

As Table 1 shows, extant literature has widely reported RSs, ISPs and ESPs as explicit support exchanges in OHCs. However, less has been done to capture implicit support topics. This group of exchanges is important, because earlier research in the broader online communities’ literature has reported the existence of off-task (e.g. indirectly health-

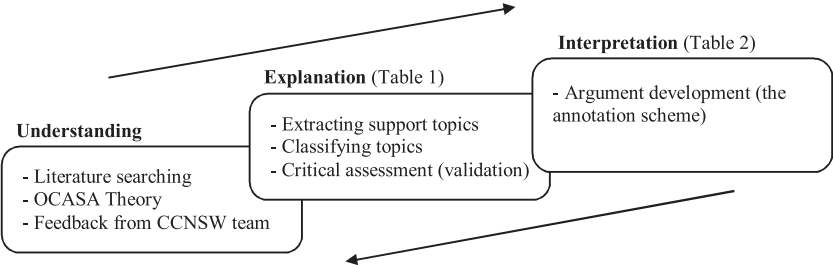


Fig. 3. Research design, influenced by a hermeneutic framework for literature review [17] and a hermeneutic framework for practical research [17].

**Table 1**  
The literature's coverage of topics and support exchange in online health communities.

Source	ISP				ESP				RS		CC			SR&D	
	ID	IT	IR	CC	E	E&S	ES	S	RIS	RES	S&C	A&A	HQ	DI	DE
Farmer et al. [20]	✓	✓	✓	✓					✓						
Bender et al. [8]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	✓
Abedin et al. [19]											✓	✓	✓	✓	
Fan, & Lederman [21]	✓	✓			✓	✓			✓	✓					
Gooden & Winefield [16]	✓	✓	✓	✓		✓								✓	
Wiljer et al. [22]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	
Gill & Whisnant [23]	✓	✓	✓	✓		✓		✓		✓				✓	✓
Fahy [18]											✓			✓	✓
Meier et al. [24]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Yan et al. [9]	✓	✓		✓	✓		✓		✓						✓
[25]	✓	✓				✓			✓	✓		✓			

related posts) content in peer-to-peer exchanges [19]. Thus, we included SR&Ds and CCs in our annotation scheme.

The conversations were then annotated using a Java web application developed specifically for this study. This application included the conversation to be annotated on the left and the annotation scheme on the right. The conversation was split into separate messages, and these in turn were split into the individual sentences to be annotated. The application randomizes the order in which conversations were given, using blocked random allocation with a block size of 10.

To test our hypotheses, independent samples t-tests were conducted using IBM-SPSS (Version 24). These tests simply compare whether the volume of activity found in initial conversations deferrers significantly between new visitors who do and do not go on to become active members of the community. To validate the annotation process, measures of inter-annotator agreement were calculated using the *agreement* module of the python package NLTK Python Package (Version 3.5).

### 3. Results

#### 3.1. Validity of the annotation process

In order to assess the consistency and subjectivity of the annotation process, two authors and the research assistant annotated 141 sentences independently. Each sentence may be annotated with any combination of the 15 leaf topics in the annotation scheme, so there are  $2^{15}$  or 32,768 distinct ways that a sentence could be annotated. This diversity of possible annotations makes it impractical to calculate a single inter-annotator agreement measure.

Instead Table 3 lists the inter-annotator agreement achieved at each level of the annotation scheme hierarchy. For example, the first row shows the agreement achieved when annotations are simplified to be one of the 16 possible combinations of the four top-level topics of the hierarchy (*self-reflection*, *requests for support*, *support provision*, and *conversational queues*) without attempting to distinguish between the lower elements of the hierarchy. The second row shows the agreement achieved when annotations are simplified to *disclosure of information*,

**Table 2**  
The annotation scheme.

Themes	Discussed topics	Descriptions	Sample quotes
Informational Support Provision (ISP)	<i>Illness (cancer) diagnosis (ID)</i>	Information about and symptoms of cancer diagnosis	<i>If it was lymphoma your white cell count would not have been normal</i>
	<i>Illness (cancer) treatment (IT)</i>	Information about treatment options, and factors to consider when making treatment decisions	<i>Without chemo the cancer would have spread more rapidly</i>
	<i>Illness (cancer) recurrence (IR)</i>	Information on procedures for regular self-monitoring, and lifestyle changes such as dietary modification, to reduce the likelihood of recurrence of cancer	<i>Stay watchful for early signs and of recurrence</i>
Emotional Support Provision (ESP)	<i>Care and coping (CC)</i>	Advice on how to cope with cancer	<i>Listen to music before chemo</i>
	<i>Encouragement (E)</i>	Support to face challenges; hopes/prayers that their situation improves	<i>Keep fighting the good fight</i>
	<i>Empathy &amp; Sympathy (E&amp;S)</i>	Expressing condolences or the ability to understand and share the feelings of another	<i>I understand how you feel; sorry to hear this news</i>
	<i>Esteem support (ES)</i>	Appreciation for the value of an individual and his or her accomplishments	<i>You inspire us with your courage</i>
	<i>Solidarity (S)</i>	Expressions of team spirit so members would not feel alone in their treatment	<i>We're all here with you</i>
Request for Support (RS)	<i>Request for informational support (RIS)</i>	Explicit request for advice or help about cancer-related questions or matters	<i>Can anyone tell me if a runny nose is a side effect to this drug &amp; how can you treat it</i>
	<i>Request for emotional support (RES)</i>	Explicit request for advice or help in coping with cancer	<i>I just need a shoulder to lean on right now</i>
Self-Reflection & Disclosure (SR&D)	<i>Disclosure of information (DI)</i>	Revealing details of the author's situation, problems faced, etc., or sharing news/events	<i>We went to the specialist today</i>
Conversational cues (CC)	<i>Disclosure of emotions (DE)</i>	Revealing details of the author's thoughts and feelings	<i>I'm completely terrified, TBH*</i>
	<i>Salutations and closings (S&amp;C)</i>	An expression of greeting or goodbye, usually at the opening or closing of the post	<i>Hello guys</i> <i>Talk later</i>
	<i>Acknowledgment &amp; Appreciation (A&amp;A)</i>	Recognizing or acknowledging the helpfulness, ideas, comments, capabilities, and experiences of other participants in the conversation	<i>Thanks for your interesting &amp; funny posts</i>
	<i>Horizontal questions (HQ)</i>	Questions which do not have a "correct" answer, but for which discussion might produce consensus or deeper understanding of the problem	<i>How did you go with the second opinion?</i>

\* TBH = to be honest.



**Table 3**  
Inter-annotator agreement at each level of the annotation scheme.

	Cohen's kappa	Krippendorff's alpha	Possible combinations
Top-level topics	0.596	0.596	16 (2 <sup>4</sup> )
Disclosure topics	0.607	0.606	4 (2 <sup>2</sup> )
Support request topics	0.832	0.831	4 (2 <sup>2</sup> )
Support provision topics	0.506	0.505	4 (2 <sup>2</sup> )
Informational support topics	0.290	0.292	16 (2 <sup>4</sup> )
Emotional support topics	0.403	0.401	16 (2 <sup>4</sup> )
Conversational queues	0.723	0.720	8 (2 <sup>3</sup> )

*disclosure of emotions*, both, or neither. The remaining rows are constructed in a similar fashion.

The table reports average pairwise Cohen's kappa [26] and Krippendorff's alpha [27]. For each measure, 1 indicates perfect agreement, while 0 indicates agreement that is no better than random chance. We also report the number of possible combinations of topics at each level of the hierarchy.

A reasonable level of agreement is achieved at most levels of the hierarchy, with the exception of support provision categories which presented a relatively poorer ability to distinguish between the different types of support topics.

### 3.2. Topics discussed in the initial conversations of new users

Table 4 lists, for each type of forum member, the proportion of conversations that contain at least one example of each of the annotation topics.

Almost all of these conversations contain self-disclosure from the new visitor. In almost all cases this includes informational disclosure, and in 43 % of cases emotional disclosure. Thus, while new visitors almost invariably describe their diagnosis and other aspects of their situation, they are less likely to discuss their thoughts and feelings openly. Roughly 40 % of new visitors make an explicit request for support, and in almost all cases, these requests focus on informational support. Explicit requests for emotional support are comparatively rare. Interestingly, many new visitors seem to join the forum with the intention of giving rather than receiving help, and tend to provide encouragement ahead of other forms of support.

Peers are quick to share the facts of their own situation when welcoming new users, presumably with the intention of showing the visitor

**Table 4**  
Proportion of conversations in which each type of forum member mentions each topic.

	new visitor	replying peers
Self-Reflection & Disclosure	93.6 %	49.3 %
Disclosure of information	92.4 %	48.9 %
Disclosure of emotions	42.7 %	20.2 %
Request for support	41.1 %	5.3 %
Request for informational support	39.6 %	5.3 %
Request for emotional support	15.8 %	0.9 %
Support Provision	36.2 %	40.9 %
Informational Support Provision	12.2 %	21.1 %
Cancer diagnosis	2.4 %	4.4 %
Cancer treatment	6.0 %	10.2 %
Cancer recurrence	1.3 %	0.2 %
Cancer care & coping	6.4 %	14.0 %
Emotional Support Provision	33.8 %	38.9 %
Encouragement	25.1 %	29.1 %
Empathy & sympathy	7.8 %	15.6 %
Esteem support	9.6 %	13.8 %
Solidarity	8.7 %	13.1 %
Conversational cues	67.6 %	45.6 %
Salutations & closings	62.2 %	43.6 %
Acknowledgment & appreciations	22.4 %	12.0 %
Horizontal questions	8.0 %	5.8 %

**Table 5**

Differences in the volume of activity related to the initial conversations of active and non-active new visitors.

	Non-active		Active		P
	Mean	(SD)	Mean	(SD)	
<b>Activity of visitor</b>					
# of sentences written by visitor	12.08	(13.27)	13.27	(14.40)	0.342
# of messages written by visitor	1.43	(0.86)	1.58	(1.18)	0.056
<b>Activity of peers</b>					
# of replying peers	0.93	(1.23)	1.12	(1.45)	0.089
# of sentences written by peers	10.72	(17.04)	14.53	(22.83)	0.017
# of messages from peers	1.23	(1.72)	1.72	(2.60)	0.004

that they are not alone but among people who share similar burdens. However, peers are less likely to discuss their thoughts and feelings about their situation, with only 20 % of conversations including any form of emotional disclosure. When providing support to new users, peers tend to focus on providing emotional support—particularly encouragement—ahead of informational support.

Influence of the volume of activity in a new visitor's conversation on their active or non-active status

Table 5 provides an analysis of whether the volume of activity in these initial conversations with new visitors differs between those who go on to become active or non-active posters. The green rows indicate that the quantity is significantly ( $P < 0.05$ ) higher in conversations started by visitors who go on to be active posters in the forum.

The volume of activity of a new visitor in their first conversation does not appear to predict whether they will become active or non-active in the forum, nor does the number of peers who contribute. However, the volume of contributions from peers (as measured by both the number of sentences and the number of messages) is significantly greater for visitors who become active, meaning that visitors are more likely to post actively if they receive a larger volume of responses from their peers in their first conversation.

Influence of topics discussed in a new visitor's initial conversation on whether they become active or non-active

Table 6 provides a breakdown of topics raised by new visitors in their initial conversations, and measures of whether the frequencies of these topics differ significantly between visitors who become active or non-active posters.

In Table 6, the green rows indicate a significant positive impact on retention, meaning that the frequency of sentences annotated against the topic is significantly higher ( $P < 0.05$ ) in conversations started by visitors who go on to stay active in the online forum. Red indicates, conversely, that the frequency is significantly lower in visitors who become active posters. It appears that new visitors who request informational support are more likely to become non-active users. Presumably, they join the forum to ask specific questions and are unlikely to actively engage after those questions are answered. In contrast, new visitors who join to help and provide support to others are more likely to remain active in the forum, particularly if they are inclined to provide emotional support in the form of encouragement or esteem support. New visitors who take care to include social niceties (salutations and closings) are also more likely to stay active in the forum. Other topics with statistically insignificant P-values represent conversation starters that do occur in the community, but which are not strongly

**Table 6**

Number of annotated sentences per conversation written by the conversation starter.

	Non-active		Active		P
	Mean	(SD)	Mean	(SD)	
<b>Self-reflection &amp; disclosure</b>	<b>8.12</b>	<b>(9.10)</b>	<b>9.31</b>	<b>(10.66)</b>	<b>0.215</b>
Disclosure of information	7.09	(8.17)	8.09	(9.32)	0.236
Disclosure of emotions	1.03	(2.04)	1.21	(2.27)	0.393
<b>Request for support</b>	<b>0.92</b>	<b>(1.43)</b>	<b>0.62</b>	<b>(1.12)</b>	<b>0.021</b>
Request for informational support	0.73	(1.13)	0.45	(0.87)	0.008
Request for emotional support	0.19	(0.50)	0.16	(0.43)	0.537
<b>Support provision</b>	<b>0.88</b>	<b>(2.00)</b>	<b>1.58</b>	<b>(3.33)</b>	<b>0.005</b>
Informational support provision	0.24	(0.85)	0.44	(1.73)	0.096
Cancer diagnosis	0.03	(0.19)	0.03	(0.24)	0.980
Cancer treatment	0.11	(0.60)	0.21	(1.03)	0.189
Cancer recurrence	0.02	(0.20)	0.04	(0.40)	0.550
Cancer care & coping	0.08	(0.37)	0.16	(0.64)	0.075
<b>Emotional support provision</b>	<b>0.64</b>	<b>(1.59)</b>	<b>1.14</b>	<b>(2.42)</b>	<b>0.008</b>
Encouragement	0.32	(0.76)	0.49	(0.78)	0.024
Empathy & sympathy	0.10	(0.40)	0.13	(0.47)	0.481
Esteem support	0.12	(0.64)	0.39	(1.88)	0.028
Solidarity	0.10	(0.38)	0.14	(0.46)	0.338
<b>Conversational cues</b>	<b>1.78</b>	<b>(2.56)</b>	<b>2.30</b>	<b>(3.07)</b>	<b>0.057</b>
Salutations & closings	1.27	(1.62)	1.72	(2.46)	0.021
Acknowledgment & appreciation	0.36	(0.99)	0.45	(0.88)	0.383
Horizontal questions	0.14	(0.79)	0.13	(0.55)	0.862

associated with a user subsequently staying active.

Table 7 provides a similar breakdown to Table 6, but focuses on the messages written by other peers in response to each new visitor.

Table 7 suggests that new visitors are more likely to become active posters if they encounter self-disclosure from other peers, and that disclosure of both information (e.g. diagnoses, experiences) and emotions (e.g. hopes, fears, concerns) is welcome. Surprisingly, it does not appear that receiving informational or emotional support strongly predicts whether new visitors stay active or become non-active, unless that support is in the form of expressions of solidarity. An additional factor is whether the new visitor is asked horizontal questions, which presumably encourage greater self-disclosure. All of these findings point towards a strong need for new visitors to quickly get a sense that they are connecting to real people who are facing similar situations. Similar to the results shown in Table 6, topics with statistically insignificant P-values represent forms of peer support that do occur in the community, but which are not strongly associated with a user subsequently staying active.

Given the above findings, Table 8 summarizes results for our hypotheses. Next section discusses the implications of these results.

**Table 7**

Number of annotated sentences per conversation written by peers responding to the conversation starter.

	Non-active		Active		P
	Mean	(SD)	Mean	(SD)	
<b>Self-reflection &amp; disclosure</b>	<b>5.56</b>	<b>(9.39)</b>	<b>9.82</b>	<b>(17.05)</b>	<b>0.001</b>
Disclosure of information	5.13	(8.68)	8.93	(15.72)	0.001
Disclosure of emotions	0.43	(1.30)	0.89	(2.43)	0.010
<b>Request for support</b>	<b>0.09</b>	<b>(0.53)</b>	<b>0.19</b>	<b>(0.75)</b>	<b>0.073</b>
Request for informational support	0.08	(0.51)	0.17	(0.62)	0.095
Request for emotional support	0.01	(0.08)	0.03	(0.25)	0.256
<b>Support provision</b>	<b>2.13</b>	<b>(4.31)</b>	<b>2.43</b>	<b>(4.06)</b>	<b>0.476</b>
Informational support provision	0.79	(2.29)	0.74	(1.71)	0.816
Cancer diagnosis	0.05	(0.34)	0.09	(0.40)	0.351
Cancer treatment	0.24	(1.16)	0.26	(0.74)	0.817
Cancer recurrence	0.00	(0.06)	0.00	(0.00)	0.460
Cancer care & coping	0.49	(1.45)	0.39	(1.47)	0.479
Emotional support provision	1.34	(2.53)	1.69	(2.84)	0.186
Encouragement	0.72	(1.73)	0.68	(1.38)	0.807
Empathy & sympathy	0.23	(0.63)	0.30	(0.87)	0.293
Esteem support	0.24	(0.80)	0.41	(1.25)	0.090
Solidarity	0.15	(0.50)	0.30	(0.79)	0.018
<b>Conversational cues</b>	<b>2.09</b>	<b>(3.77)</b>	<b>2.45</b>	<b>(4.19)</b>	<b>0.361</b>
Salutations & closings	1.78	(3.01)	1.99	(3.48)	0.502
Acknowledgment & appreciation	0.25	(1.07)	0.23	(0.67)	0.846
Horizontal questions	0.07	(0.31)	0.23	(1.31)	0.046

## 4. Discussion and implications

### 4.1. Primary findings

While previous researchers have explored support exchange in the OHC context (e.g. [28,29]) and stressed the importance of newcomers on the community dynamics (e.g. [12] & [11]), no prior study specifically focused on the impact of new members' initial posts and the subsequent effects on their activity level. Thus, we firstly sought to establish what topics were discussed in the initial conversations of new users. Our results showed that almost all of these conversations contained self-disclosure from the new visitors. Informational disclosure occurred in almost all cases, and emotional disclosure in almost half of cases. Smithson et al. [13] reported similar results, and highlighted that OHC members often initiate a thread by asking for 'advice' that they could either find for themselves easily online or that had an obvious 'safe' answer. Furthermore, we found that almost half of new visitors make an explicit request for support, and in almost all cases these requests are for informational support, rarely for emotional support. Emotional support and disclosure are important, as earlier research has stressed on the deeply remedial potential of emotional reciprocity [30]. Many new visitors seem to join the forum with the intention of giving rather than receiving help, and tend to provide encouragement ahead of other forms of support. This is in line with Butler et al.'s (2014) characterisation of bond-based communities as a space in which members act altruistically for the good of the community. Peers are quick to share the facts of their own situation when welcoming new users, presumably with the intention of showing the visitor that they are not alone. However, peers are less likely to discuss their thoughts and feelings about their situation, with only 20 % of conversations including any form of emotional disclosure.

Next, the first hypothesis examined new visitor's conversation and whether it influences their active or non-active status. We found a partial support for this hypothesis. H1a was not supported as the volume of activity of a new visitor in their first conversation was not associated with becoming an active or non-active user in the forum. In contrast, H1b was supported as the topics of new members' conversation either positively or negatively influenced on their activity level. In particular, we found that new visitors who request informational support are more likely to become non-active users. Presumably, they join the forum to ask specific questions and are unlikely to actively engage after those questions are answered. In contrast, new visitors who join to help and provide support to others are more likely to remain active in the forum, particularly if they are inclined to provide emotional support in the form of encouragement or esteem support. This is in line with some other studies: Smithson et al. [13] reported that the offering of informational support is typically hedged with expressions of emotional support, and that a key function of the activity on OHCs could be seen as connecting, developing relationships, or 'just being there' for someone who is struggling. New visitors who take care to include social niceties (salutations and closings) are also more likely to stay active in the forum.

In regards to our second hypothesis, we found support for H2a as findings showed that new visitors are more likely to become active posters if they encounter self-disclosure from other peers, and that disclosure of both information (e.g. diagnoses, experiences) and emotions (e.g. hopes, fears, concerns) is welcome. Surprisingly, receiving informational or emotional support is not associated with new visitors staying active, unless that support is in the form of expressions of solidarity. An additional factor is whether the new visitor is asked horizontal questions, which presumably encourage greater self-disclosure. Furthermore, H2b was accepted as the volume of contributions from peers is significantly associated with becoming active, meaning that visitors are more likely to actively post if they receive a larger volume of responses from their peers in their first conversation. However the number of peers who contributed to initial posts of newcomers

**Table 8**  
Summary of hypotheses outcomes.

Hypothesis	Outcome
H1a: Topic of new members' initial post influences their activity level	Not supported
H1b: Volume of new members' initial post influences their activity level	Supported
H2a: Topic of community reaction to a new member's initial posts influences the member's activity level	Supported
H2b: Volume of community reaction to a new member's initial posts influences the member's activity level	Supported

presented no impact on the activity level. All of these findings point towards a strong need for new visitors to quickly get a sense that they are connecting to real people who are facing similar situations.

#### 4.2. Theoretical implications

Butler et al. OCASA theory is increasingly receiving interest from OHCs scholars for examining community design and members' social interactions (e.g. [31–33]). However, more work is still needed to extend this into the bond-based online communities as Butler et al. concede “the extent to which such other factors drive member expectations and participation decisions in ways that are inconsistent with cost–benefit logic is a limit on the generalizability of the simulation model presented here” ([14], page 723). Motivated by this, the present study has contributed to the OCASA theory by examining new OHC members' initial posts and community reactions to them, and extending the theory to a bond-based context (i.e. OHCs) in which members do not have a financial cost–benefit motivation.

As Fig. 4 summarizes, our results have implications for the three stages of the theory. We found that initial posts made by new members, coupled with the community reactions to these posts, can determine active or non-active status of individuals. This is in line with the OCASA, as initial posts reflect new members' expression of expectations from the community, and the subsequent reactions to these posts can determine whether the members may actively engage with the community. Given the support for H1b, the content of messages and the type of support exchanged as a result of initial posts can predict new members' activity levels. This is in-line with Wang et al. [34] findings that emotional support is correlated with length of membership. Furthermore, the support for H2 shows that subsequent to receiving support from the community, new members are more likely to stay active in the community. As also echoed by Introne et al. [31], this perception gets reinforced among new members that the forum is a good fit for their needs, which then may encourage them to stay in the community.

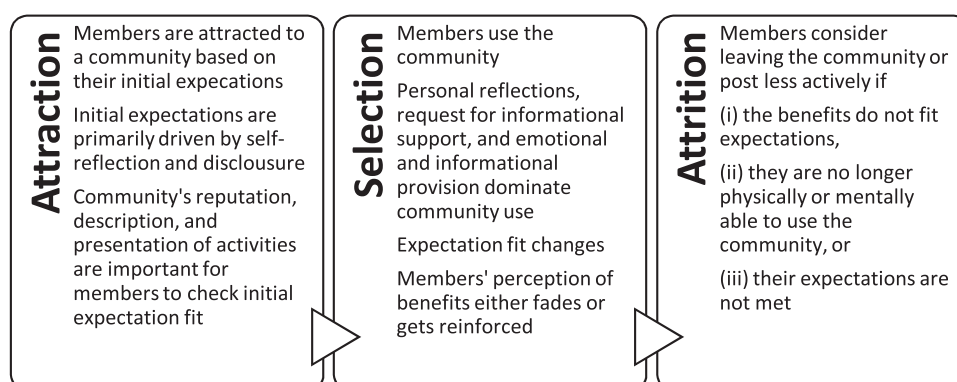
This study highlighted that OHC members engage in reciprocal exchanges of self-disclosure coupled with provision of information and emotional support as well as social niceties and solidarity. We found that these exchanges are often a way for members to express the desire to connect with others and develop relationships. Smithson et al. [13] produced a similar finding—that the same OHC members who post

conventional advice on health care often initiate new posts asking for advice in similar situations. These findings extend and contribute to earlier studies by Yan et al. [9] and others that reported personal benefits (i.e. reputation enhancement, perceived social support, and a sense of self-worth) and avoidance of costs (i.e. executional and cognitive costs) are key drivers of memberships in OHCs.

#### 4.3. Practical implications

In addition to the theoretical contributions, this study also offers support and sheds light on OHCs practical applications. Firstly, OHCs managers and owners should explicitly highlight on their website the benefits of the online community for patients and other health users. While OHCs offer valuable health related informational support, community managers should use the ‘About Us’ or other community introduction pages to stress the importance of connecting with like-minded peers and exchanging emotional support via an OHC. This, as per OCASA, helps to ensure new members' expectations of the community management activities and initiatives are realized. Secondly, our results inform health organisations' policymaking and advocacy efforts towards strengthening online support for cancer patients and other health users (e.g. carers). The results of this study encourage health organisations to take advantage of online health communities to supplement their existing health support services. For instance, Cancer Council NSW is currently hosting one of the most active and vibrant online health forums in NSW which is increasingly attracting people from around the country and overseas. This is because its cancer health advice and peer support can be used by patients across the globe.

Lastly, as Yan et al. [9] have also reported, OHCs managers and owners should encourage self-disclosure and personal reflections and publicize the positive experiences and impacts that this may have for both the persons who actively share (active posters) and others. Other studies have also found that publicizing results of peer support plays an important role in offering hope, facilitating coping, and enhancing quality of life (e.g. [2,12]). For instance, research shows that patients may benefit from being exposed to others who have recovered to some degree from the same disorder [35].



**Fig. 4.** The OCASA theory for bond-based communities.

## 5. Conclusions and recommendations for future research

In this paper, we examined whether the content and volume of new OHC members' initial posts is associated with their subsequent activity levels. Findings revealed that the initial posts of new OHC members are strongly associated with whether members actively post to the community following their initial posts and the community's reactions to these posts. Moreover, in line with previous studies, this research revealed high proportions of peer exchanges on Request for Support and Support Provision topics. We also documented numerous exchanges relating to Self-Reflection & Disclosure and Conversational Cues in the online cancer community under study. Findings revealed that users are more likely to exchange information support than emotional support. Many new users reflect on their personal circumstances and use initial conversations as a mechanism to connect to others and feel part of the community.

We extended OCASA theory to a bond-based context, and presented members' expectation from the community through topics of their initial posts, and how these topics and their associated volumes can predict members' selection and attrition. However, our contribution to OCASA stays limited to the assessment of members' initial posts, which creates opportunities for future studies to examine pattern of activities and the support exchanged between members over a period of time, and relate them to members' first post profile. Some like Panciera et al. [36] suggest that active members of online communities like Wikipedia are born and not made, and thus future research could examine such hypothesis in the OHCs context.

Our findings also imply that OHC members do not simply leave the community if their expectations are not satisfied in the 'selection' stage. Rather, they may swing between active and non-active status, meaning active posters may temporarily engage in the community passively. However, one limitation of this study was the absence of page view measurement, so we were unable to track passive use and abandonments. Consequently, we were unable to test the validity of this implication of our findings since we solely relied on textual data; other researchers could collect page view data to do so. Future studies can also extend the definition of non-active/active representation using other methods (e.g. tracking or interviewing users) to make a clearer distinction between activity levels and users' needs and expectations in OHCs.

Furthermore, future researchers should investigate the behavior of other cohorts of OHC users, particularly carers of patients. Empirical research is also needed to link OHC members' activity levels with their health conditions, living environments, cancer types and other factors. Finally, researchers are encouraged to study the role of moderation in attraction, selection, and attrition of users in OHCs. While some studies recommend a slow and infrequent style of moderator intervention [37], others argue for frequent and personalized moderator engagement [38,39].

### Author's statement

We confirm that all authors have contributed significantly to the manuscript, with the first author leading the project and developing theoretical underpinnings, the second author developing the annotation tool and statistical tests, and the third author developing research background and support. All authors have seen and approved the manuscript. We confirm that informed consent was obtained ethics approval for conducting this study from University of Technology Sydney. The privacy rights of human subjects and anonymity of participants have been maintained.

We also confirm that this is an original work that is not currently under review or publication consideration at any other outlet

What we already know on this topic:

- Use of online health communities by health users and practitioners is on the rise
- These communities are effective environments for peer support and information exchange
- Usefulness of these communities depends on active engagement of community members
- It is unclear what exact topics are discussed in different online health communities and why activity levels rise or drop

What this study added to knowledge:

- Volume of new members' initial posts can predict future activity levels (i.e. active vs. non-active)
- Community reactions to initial posts can predict future activity levels (i.e. active vs. non-active)

- Key topics exchanged in online health communities are: informational support provision, emotional support provision, request for help, self-reflection & disclosures, and conversational cues
- Members of the community overwhelmingly disclose personal reflections to bond with the community
- Many members actively post to the community solely to provide informational and emotional support to others

### Declaration of Competing Interest

None.

### Acknowledgements

This paper was produced as part of a grant funded by Cancer Council NSW (PRO18-6515). Authors would like to thank Annette Beattie, Kate Fenerty, and Rhiannon Davies from CCNSW for their inputs and support.

### Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.ijmedinf.2020.104216>.

### References

- [1] L. Chen, A. Baird, D. Straub, An analysis of the evolving intellectual structure of health information systems research in the information systems discipline, *J. Assoc. Inf. Syst.* 20 (8) (2019) 1023–1074.
- [2] B. Abedin, S. Erfani, D. Milne, A. Beattie, K. Fenerty, Unpacking support types in online health communities: an application of attraction-selection-attrition theory, *Pacific Asia Conference on Information Systems*, China, 2019.
- [3] S.S. Erfani, B. Abedin, Impacts of the use of social network sites on users' psychological well-being: a systematic review, *J. Assoc. Inf. Sci. Technol.* 69 (7) (2018) 900–912.
- [4] J. Zhou, M. Zuo, C. Ye, Understanding the factors influencing health professionals' online voluntary behaviors: evidence from YiXinLi, a Chinese online health community for mental health, *Int. J. Med. Inform.* 130 (2019) 103939.
- [5] H. Wu, N. Lu, Online written consultation, telephone consultation and offline appointment: an examination of the channel effect in online health communities, *Int. J. Med. Inform.* 107 (2017) 107–119.
- [6] X. Liu, M. Sun, J. Li, Research on gender differences in online health communities, *Int. J. Med. Inform.* 111 (2018) 172–181.
- [7] Y. Li, Y. Song, W. Zhao, X. Guo, X. Ju, D. Vogel, Exploring the Role of Online Health Community Information in Patients' Decisions to Switch from Online to Offline Medical Services, *Int. J. Med. Inform.* 130 (2019) 103951.
- [8] J.L. Bender, M.C. Jimenez-Marroquin, A.R. Jadad, Seeking support on Facebook: a content analysis of breast cancer groups, *J. Med. Internet Res.* 13 (1) (2011) 16.
- [9] Z. Yan, T. Wang, Y. Chen, H. Zhang, Knowledge sharing in online health communities: a social exchange theory perspective, *Inf. Manag.* 53 (5) (2016) 643–653.
- [10] M. Burke, R. Kraut, E. Joyce, Membership claims and requests: conversation-level newcomer socialization strategies in online groups, *Small Group Res.* 41 (1) (2010) 4–40.
- [11] B. McInnis, G. Leshed, D. Cosley, Crafting policy discussion prompts as a task for newcomers, *Proc. ACM Hum. Interact.* 2 (CSCW) (2018) 1–23.



- [12] J. Levine, H.S. Choi, R.L. Moreland, Newcomer influence and creativity in work groups, *The Oxford Handbook of Group Creativity and Innovation*, (2019), p. 51.
- [13] J. Smithson, S. Sharkey, E. Hewis, R. Jones, T. Emmens, T. Ford, C. Owens, Problem presentation and responses on an online forum for young people who self-harm, *Discourse Stud.* 13 (4) (2011) 487–501.
- [14] B.S. Butler, P.J. Bateman, P.H. Gray, E.I. Diamant, An attraction-selection-attrition theory of online community size and resilience, *Mis Q.* 38 (3) (2014) 699–728.
- [15] S.S. Erfani, B. Abedin, F. Daneshgar, A qualitative evaluation of communication in ovarian cancer facebook communities, *IEEE Conference on Information Society (I-Society)* (2013) 270–272.
- [16] R.J. Gooden, H.R. Winefield, Breast and prostate cancer online discussion boards: a thematic analysis of gender differences and similarities, *J. Health Psychol.* 12 (1) (2007) 103–114.
- [17] S.K. Boell, D. Cecez-Kecmanovic, A hermeneutic approach for conducting literature reviews and literature searches, *Commun. Assoc. Information Syst.* 34 (12) (2014) 257–286.
- [18] P.J. Fahy, Addressing some common problems in transcript analysis, *Int. Rev. Res. Open Distrib. Learn.* 1 (2) (2001) 1–6.
- [19] B. Abedin, F. Daneshgar, J. D'Ambra, Pattern of non-task interactions in asynchronous computer-supported collaborative learning courses, *Interact. Learn. Environ.* 22 (1) (2014) 18–34.
- [20] A. Farmer, C. Holt, M. Cook, Social networking sites: a novel portal for communication, *Postgraduate Med J.* 85 (2009) 455–459.
- [21] H. Fan, R. Lederman, Online health communities: how do community members build the trust required to adopt information and form close relationships? *Eur. J. Inf. Syst.* 27 (1) (2018) 62–89.
- [22] D. Wiljer, S. Urowitz, L. Barbera, M.L. Chivers, N.K. Quartey, S.E. Ferguson, et al., A qualitative study of an internet-based support group for women with sexual distress due to gynecologic cancer, *J. Cancer Educ.* 26 (3) (2011) 451–458.
- [23] P.S. Gill, B. Whisnant, A qualitative assessment of an online support community for ovarian cancer patients, *Patient Relat. Outcome Meas.* 3 (2012) 51–58.
- [24] A. Meier, E. Lyons, G. Frydman, M. Forlenza, B. Rimer, How cancer survivors provide support on cancer-related Internet mailing lists, *J. Med. Internet Res.* 9 (2) (2007) e12.
- [25] Q. Zheng, T. Yao, X. Fan, Improving customer well-being through two-way online social support, *J. Serv. Theory Pract.* 26 (2) (2016) 179–202.
- [26] J. Cohen, A coefficient of agreement for nominal scales, *Educ. Psychol. Meas.* 20 (1) (1960) 37–46.
- [27] K. Krippendorff, Estimating the reliability, systematic error and random error of interval data, *Educ. Psychol. Meas.* 30 (1) (1970) 61–70.
- [28] X. Wang, K. Zhao, N. Street, Analyzing and predicting user participations in online health communities: a social support perspective, *J. Med. Internet Res.* 19 (4) (2017) e130.
- [29] M. Zhang, C.C. Yang, Using content and network analysis to understand the social support exchange patterns and user behaviors of an online smoking cessation intervention program, *J. Assoc. Inf. Sci. Technol.* 66 (3) (2015) 564–575.
- [30] G. Jefferson, J.R.E. Lee, The rejection of advice: managing the problematic convergence of a 'troubles-telling' and a 'service encounter', *J. Pragmat.* 5 (1981) 399–422.
- [31] J. Introne, I. Erickson, B. Semaan, S. Goggins, Designing sustainable online support: examining the effects of design change in 49 online health support communities, *J. Assoc. Inf. Sci. Technol.* 71 (4) (2019) 379–394.
- [32] R. Kumi, R. Sabherwal, Knowledge sharing behavior in online discussion communities: examining behavior motivation from social and individual perspectives, *Knowl. Process. Manag.* 26 (2) (2019) 110–122.
- [33] J. Introne, B. Semaan, S. Goggins, A sociotechnical mechanism for online support provision, *CHI Conference on Human Factors in Computing Systems* (2016) 3559–3571.
- [34] Y.C. Wang, R. Kraut, J.M. Levine, To stay or leave? The relationship of emotional and informational support to commitment in online health support groups, February, *ACM 2012 Conference on Computer Supported Cooperative Work* (2012) 833–842.
- [35] B. Melling, T. Houguet-Pincham, Online peer support for individuals with depression: a summary of current research and future considerations, *Psychiatr. Rehabil. J.* 34 (3) (2011) 252.
- [36] K. Panciera, A. Halfaker, L. Terveen, Wikipedians are born, not made: a study of power editors on wikipedia, *ACM International Conference on Supporting Group Work (GROUP'09)*, ACM, New York, NY, USA, 2009, pp. 51–60.
- [37] K. Wise, B. Hamman, K. Thorson, Moderation, response rate, and message interactivity: features of online communities and their effects on intent to participate, *J. Comput. Commun.* 12 (1) (2006) 24–41.
- [38] Y. Ren, R.E. Kraut, Agent-based modeling to inform online community design: impact of topical breadth, message volume, and discussion moderation on member commitment and contribution, *Hum.-Comp. Interact.* 29 (4) (2014) 351–389.
- [39] D. Priharsari, B. Abedin, E. Mastio, Value co-creation in firm sponsored online communities: what enables, constrains, and shapes value, *Internet Res.* (2020), <https://doi.org/10.1108/INTR-05-2019-0205>.