

Shining a light:
Active participation in a mental health
Internet support group

by

Bradley John Carron-Arthur

A thesis submitted for the degree of

Doctor of Philosophy

of The Australian National University

August 2017

© Copyright by Bradley John Carron-Arthur 2017

All Rights Reserved

Thesis by compilation

This piece of work comprises a 'thesis by compilation' as permitted by The Australian National University's *Research Awards Rules*. According to these rules:

A thesis by compilation includes papers (e.g. articles, chapters) where the student is the sole or joint author that are in the process of being prepared or approved for publication, have been accepted for publication, or have been published.

A thesis by compilation contains

- An introduction to the field of study and the hypothesis or research questions, how these are addressed through the ensuing chapters, and a general account of the theory and methodological components of the research where these components may be distributed across separate papers/chapters.
- Linking text to establish the relationship between one chapter and the next such as through a foreword to each chapter.
- A conclusion drawing together the published papers or works in a cohesive manner, and addresses how the individual publications link to the theory and methodology adopted and evaluate the contribution that the research in the submitted publications makes to the advancement of the research area.

A thesis by compilation consists primarily of published papers and manuscripts accepted for publication, and not primarily of manuscripts submitted and under review by referees or manuscripts under revision following referees reports.

The scope and quality of a thesis by compilation is commensurate with the contribution to knowledge expected of a candidate for the relevant degree.

A thesis by compilation includes a signed declaration that specifies:

- Title, authorship and publication outlet of each paper.
- The current status of each paper (In press, Accepted, Under review, In preparation).
- The extent of the contribution of the candidate to the research and the authorship of each paper.

Declaration

I declare that the work contained in this thesis is the result of original research and has not been submitted to any other University or Institution.



19 August 2017

Bradley John Carron-Arthur

Date

Paper 1

Carron-Arthur, B, Ali, K, Cunningham, JA, Griffiths, KM. From help-seekers to influential users: a systematic review of participation styles in online health communities. *Journal of Medical Internet Research*. 2015; 17(12): e271.

Paper 2

Carron-Arthur, B, Cunningham, JA, Griffiths, KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions*. 2014; 1(4): 165-8.

Paper 3

Griffiths, KM, Carron-Arthur, B, Reynolds, J, Bennett, K, Bennett, A. User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study. *Internet Interventions*. 2017; 7: 9-15.

Paper 4

Carron-Arthur, B, Reynolds, J, Bennett, K, Bennett, A, Cunningham, JA, Griffiths, KM. Community structure of a mental health Internet support group: Modularity in user thread participation. *JMIR Mental Health*. 2016; 3(2): e20.

Paper 5

Carron-Arthur, B, Reynolds, J, Bennett, K, Bennett, A, Griffiths, KM. What's all the talk about? Topic modelling in a mental health Internet support group. *BMC psychiatry*. 2016; 16(1): 367.

Paper 6

Carron-Arthur, B, Cooper, M, Kazan, D, Ali, K, Griffiths KM. Are first impressions predictive of future engagement? Qualitative analysis and blinded judgements of users' initial post in a mental health Internet support group. In Preparation.


Table 1 contains a summary of the contribution I made to each paper in terms of (i) the study concept and design, (ii) analysis and interpretation and (iii) drafting and revising the manuscript.

Table 1. Author contributions: Bradley Carron-Arthur

Paper	Study concept and design	Analysis and interpretation	Drafting and revising	Average
1	80%	90%	90%	87%
2	65%	80%	90%	78%
3	10%	20%	5%	12%
4	65%	80%	90%	78%
5	85%	90%	80%	85%
6	65%	85%	90%	80%

Signature of senior author

I agree that Bradley John Carron-Arthur made the contributions outlined above to the papers on which I am the senior author.



19 August 2017

Emeritus Professor Kathleen Margaret Griffiths

Date

Copyright

Paper	Terms of copyright
1	©Bradley Carron-Arthur, Kathina Ali, John Alastair Cunningham, Kathleen Margaret Griffiths. Originally published in the <i>Journal of Medical Internet Research</i> (http://www.jmir.org), 01.12.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the <i>Journal of Medical Internet Research</i> , is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/ , as well as this copyright and license information must be included.
2	© Bradley Carron-Arthur, John Alastair Cunningham, Kathleen Margaret Griffiths. Originally published in <i>Internet Interventions</i> (https://doi.org/10.1016/j.invent.2014.09.003). This is an open-access article distributed under the terms of the Creative Commons Attribution – Non Commercial – No Derivs (CC BY-NC-ND) licence: for non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.
3	©, Kathleen Margaret Griffiths. Bradley Carron-Arthur Julia Reynolds, Kylie Bennett, Anthony Bennett. Originally published in <i>Internet Interventions</i> (https://doi.org/10.1016/j.invent.2016.11.003). This is an open-access article distributed under the terms of the Creative Commons Attribution – Non Commercial – No Derivs (CC BY-NC-ND) licence: for non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.
4	©Bradley Carron-Arthur, Julia Reynolds, Kylie Bennett, Anthony Bennett, John Alastair Cunningham, Kathleen Margaret Griffiths. Originally published in <i>JMIR Mental Health</i> (http://mental.jmir.org), 30.05.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in <i>JMIR Mental Health</i> , is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/ , as well as this copyright and license information must be included.
5	© Bradley Carron-Arthur, Julia Reynolds, Kylie Bennett, Anthony Bennett, Kathleen Margaret Griffiths. Originally published in <i>BMC Psychiatry</i> (https://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-016-1073-5). No changes have been made. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

Acknowledgements

Foremost, thank you to the BlueBoard community. Without you this research would not have been possible. I hope this thesis sheds some light on those who shine a light for others.

I would like to thank the BlueBoard moderators, Jan, Jo, Michelle and Cathy, and particularly Julia Reynolds who led the team as the Clinical Services Manager, for all the insights into the BlueBoard community that you provided me.

Thank you to the e-hub team, including Kylie and Anthony Bennet for making BlueBoard a reality.

Thank you to the Australian Government for providing the funding for BlueBoard and for my Australian Postgraduate Award Scholarship.

Thank you to Professor John Cunningham for your supervision, particularly in the early days of this thesis when you got me up and moving quickly and in the direction of thesis by publication. Thank you to Dr Michelle Banfield for your support throughout and your diligent work in reviewing the final product.

Thanks to all my friends and colleagues at the Centre for Mental Health Research and the Research School of Population Health who've made this journey so enjoyable. In particular, a special thank you to Fiona for wise words, deep chats, and for generally looking out for me, both as a friend and a colleague. Special thanks also to Kathina, Dom, Brendan and Marita for helping me out with studies, even when you were flat-out busy yourselves.

Thank you Marion Stanton for being such a great mentor in all the co-curricular that rounded out my PhD journey and for nominating me for Postgraduate Student of the Year. I never saw it coming.

Thank you to all my friends and family outside of ANU, especially my fiancé Sarah for buoying me through some big ups and downs over the last few years. But really this journey started long ago and I particularly have my mum, Dr Philippa Carron, to thank for making my education and wellbeing such a high priority. Dad, you've been an inspiration too.

Last but not least, Professor Kathy Griffiths, how can I thank you enough? I've developed so much under your mentorship and supervision. I feel like a different person. Thank you immensely.

Abstract

Internet Support Groups (ISGs) are a valued and popular source of health information and support among consumers and carers. Although ISGs are premised upon mutual help, it has been observed that only a small minority of users, of the order of 1%, are responsible for the majority of activity. Despite their potential importance to the outcomes and sustainability of online groups, little is known about the characteristics of these participants or the nature of their participation.

This thesis comprises a systematic review of the literature on styles of participation in ISGs followed by a series of five empirical studies focusing on the nature of participation in a Mental Health Internet Support Group (MHISG). These studies sought to address fundamental gaps in our knowledge regarding active participation in an MHISG, posing the questions: 'Who participates?', 'With whom do they communicate?', 'What do they communicate about?' and 'How do these factors differ as a function of user engagement?'. These questions were addressed using log data generated by all active users (n=2932) of the MHISG 'BlueBoard' and a mixture of qualitative and quantitative methods including novel analyses, such as social network modularity and topic modelling algorithms.

It was found that the demographic characteristics of higher- and lower-engaged users were broadly similar, although the members of the higher-engaged group were older and more likely to identify as consumers. Network analysis demonstrated users communicated with each other in a pattern that resembled five generational cohorts transcending disorder-specific sub-forums, in which the highest-engaged users of each cohort were central and registered earlier than the majority of other users. Topic modelling and qualitative content analysis revealed the content of the communications of the two groups differed. The communications of higher-engaged users appeared to reflect a consumer model of recovery and those of lower-engaged users a medical model of recovery. However, higher-engaged users modified the content of their responses when communicating with lower-engaged users. Qualitative analysis of users' initial posts revealed higher- and lower-engaged users differed in terms of their 'awareness' characteristics at the outset of participation, with higher-engaged users demonstrating greater interpersonal-, mental health- and self-awareness.

Based on these findings, this thesis presents ‘The Tripartite Model of MHISG Participation’ which, contrary to prevailing assumptions, posits that differences in posting frequency are associated with different styles of active participation across the spectrum of engagement. The higher end comprises a minority group of users—referred to as ‘mutual helpers’—who are central, aware and proactive about participating in peer support for their ongoing recovery. At the lower end, the majority of users, referred to as ‘active help seekers’ and ‘active help providers’, participate in transient and asymmetrical exchanges, often with ‘mutual helpers’. Those who do not post are ‘passive followers and help seekers’. The model is iterated for each cohort. In addition to extending our scientific knowledge base, and informing the above new model of user participation, these findings are of potential relevance to the design of future research studies, managers of Internet support groups and policy makers.

Table of Contents

Thesis by compilation	iii
Declaration.....	iv
Copyright.....	vi
Acknowledgements.....	vii
Abstract.....	ix
Chapter 1.....	1
1.1 Overview	1
1.1.1 Rationale	1
1.1.2 Method	2
1.1.3 Thesis structure.....	3
1.2 Background literature and policy.....	3
1.2.1 Mental illness	3
1.2.2 E-mental health.....	6
1.2.3 Peer support.....	8
1.2.4 Peer support on the Internet.....	10
1.2.5 Mental health Internet support groups.....	11
1.2.6 Active participation in mental health Internet support groups.....	14
1.3 The current research.....	16
1.3.1 Aims and overview of the studies in the thesis	16
1.3.2 BlueBoard.....	19
1.4 References	21
Chapter 2.....	27
Foreword.....	27
From help-seekers to influential users: A systematic review of participation styles in online health communities	29
Chapter 3.....	45
Foreword.....	45
Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law.....	46
Chapter 4.....	51
Foreword.....	51
User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study.....	52
Chapter 5.....	59
Foreword.....	59

Community structure of a mental health Internet support group: Modularity in user thread participation	61
Chapter 6	71
Foreword	71
What's all the talk about? Topic modelling in a mental health Internet support group.....	72
Chapter 7	85
Foreword	85
Are first impressions predictive of future engagement? Qualitative analysis and blinded judgements of users' initial post in a mental health Internet support group.....	87
Chapter 8	117
8.1 Summary of research findings	117
8.1.1 In what ways has participation previously been defined and measured in an MHISG, and other online health communities? (RQ1)	117
8.1.2 What is the nature of active participation in BlueBoard (in terms of user characteristics, activity, connectivity and content) and does it differ between differentially-engaged users? (RQs 2-6)	118
8.2 Findings and methodology in the context of other research	119
8.3 Synthesis: A model of participation.....	122
8.4 Implications for practice and policy	126
8.4.1 Presentation of the MHISG to new users	126
8.4.2 Supporting active help providers to support others	130
8.5 Methodological strengths and limitations of the investigation	130
8.5.1 Strengths.....	130
8.5.2 Limitations	131
8.6 Future research	133
8.6.1 Direct extensions	133
8.6.2 Broader implications for future research	134
8.7 Conclusion	135
8.8 References.....	136

Chapter 1

1.1 Overview

This chapter provides context for the compilation of publications included in this thesis. It outlines the rationale underpinning and linking these publications and contextualises them with respect to each other and in the broader relevant literature. Further context is provided in the forewords to each individual publication (Chapters 2-7) and in the concluding chapter of this thesis (Chapter 8).

1.1.1 Rationale

This thesis comprises an investigation into the nature of active participation in a Mental Health Internet Support Group (MHISG). Although there is a high prevalence of participation in such groups around the world [1, 2] and extensive interest in online health behaviour more broadly [3], there is little research on this topic [4]. Arguably, the resulting gap in knowledge about a fundamental aspect of MHISGs, and the consequent reliance on assumptions lacking empirical validation, has limited progress in what constitutes a promising area for mental health care [5-8].

Currently, MHISGs are represented in the literature as an homogenous intervention comprising peer-to-peer mutual support [9, 10]. *Active participation* in this intervention is defined by the act of contributing content to the MHISG. Thus, active participation is typically measured by the basic unit of contribution, a post, and the extent of *engagement* in the intervention by an individual is typically defined by their posting frequency [11]. Recent research, including a replication study reported in the current thesis, has found that the distribution of engagement in an MHISG follows an inverse power law, in which the majority of the content (of the order of 75%) is created by a small minority of users (1%) [12-14]. This phenomenon is referred to as participation inequality in the broader literature on online communities [15]. Our systematic review of studies of active participation in health ISGs concluded that such research is typically underpinned by the assumption that more highly-engaged users receive a greater dose of the intervention than low-engaged users but that the mode of action of the intervention is the same for each group [11]. However, the validity of

this assumption and whether there are meaningful differences between users who are differentially engaged in MHISGs had not yet been investigated.

The research question guiding the empirical studies of this thesis was therefore ‘are there systematic differences in the nature of participation between users at higher and lower levels of engagement?’. Given that the nature of participation is multifaceted, this overarching research question is asked repeatedly in multiple studies that investigate various fundamental aspects of active participation including ‘what are the characteristics of people who participate?’, ‘with whom do they communicate?’, ‘what do they communicate about?’.

The answer to these research questions has important implications. An understanding of the nature of active participation in an MHISG is necessary if we are to understand the mechanisms that might underpin the effectiveness of these groups. From a community management perspective, such knowledge is required to inform the optimal development, management and service delivery of MHISGs. From a research perspective, such information has the potential to affect the design and measures employed in future studies of the effects of MHISGs should the findings suggest that active participation is heterogeneous rather than homogenous in nature. From a policy perspective, determining how MHISGs are used can inform their potential role within the mental healthcare system and what aspects require investment. The existence of people who voluntarily provide support to thousands of people—an activity which may be viewed by and thus indirectly support thousands more—highlights an area that warrants further investigation to inform what role these people, and the communities to whom they are central, are best placed to perform.

1.1.2 Method

Previous studies on the nature of MHISG participation have been limited with respect to their aims and methods [4]. Few studies have employed analyses other than overall descriptive summaries of user characteristics and content. Almost all research has relied upon surveys and manual content analyses. In contrast, the studies included in this thesis address the above research questions using log data and information collected during user registration. These studies employ innovative analyses, such as a network modularity algorithm and a topic modelling algorithm, to provide new and informative perspectives on the nature of participation in an MHISG, particularly how it differs between differentially engaged users. The current approach thus comprises analyses that investigate key questions systematically, free

from selection effects associated with the administration of surveys, and in the case of the topic model less subject to the biases associated with manual content analyses. In doing so, this thesis sought to provide data that would inform a model of the nature of participation in an MHISG.

All empirical studies were conducted on a single MHISG, BlueBoard. Background information regarding BlueBoard is provided at the end of this chapter.

1.1.3 Thesis structure

This thesis comprises eight chapters. Chapter 1 (this chapter) provides a context for the ensuing compilation of papers, presenting brief overviews of the relevant published research and policy literature and a description of the rationale of the studies included in the thesis, along with the specific research questions addressed in each study. Chapter 2 provides a systematic review of the empirical literature on participation in an MHISG (Chapter 2). This is followed by the empirical findings of the thesis in the form of four published papers (Chapters 3 to 6) and a yet-to-be-submitted paper (Chapter 7). The thesis concludes with a chapter that summarises the findings, synthesises them into a model of participation and discusses the possible implications in the context of extant literature and policy (Chapter 8).

1.2 Background literature and policy

This section includes a brief overview of literature relevant to the ensuing chapters outlining: 1) key issues in population mental health; 2) advantages of e-mental health services in addressing these issues; 3) definitions and underlying theory regarding peer support; 4) the extent of participation in online peer support; 5) findings from studies of mental health Internet support groups; and 6) the current conceptualisation of active participation in an MHISG.

1.2.1 Mental illness

Mental illness is a significant global health issue. Common mental disorders such as anxiety and depression are highly prevalent throughout the world [16, 17]. Inadequate access to

mental healthcare services is similarly a common issue, even among high-income countries [18]. The strategies for addressing mental health issues are broadly similar across countries [19]. In view of these similarities, and since the majority of participants in the studies included in this thesis are Australian residents, the current review focuses primarily on literature and data from an Australian context.

1.2.1.1 Prevalence. In Australia, the one-year prevalence of mental disorders among people aged 16–85 years is 20% [20]. Over a lifetime the prevalence is 45.5% [20]. Of all major health conditions in Australia, mental illness is responsible for the third largest burden of disease (12%), behind cancer (19%) and cardiovascular diseases (15%) [21]. The life expectancy of Australians with severe mental disorders is between 10 to 36 years lower than the population average [22]. The annual cost of mental illness in Australia, including the cost of treatment and lost productivity, has been estimated to be \$20 billion [23].

The symptoms, chronicity and severity of mental illness can differ substantially between and within different mental health conditions. The most common disorders in Australia are anxiety disorders (14.4%), affective disorders (6.2%), and substance use disorders (5.1%) [20]. Mental disorders are often experienced in combination. Of the 20% of Australians who experience a mental disorder in any one year, 57.5% have one disorder and 42.5% have two or more disorders [20]. Based on DSM-IV and ICD-10 criteria, one-fifth (20.5%) are classified as severe, one third (33.2%) are classified as moderate, and just under half (46.3%) are classified as mild [24]. A higher proportion of people with affective disorders experience a severe mental disorder than people with an anxiety disorder or substance use disorder [24]. The reported prevalence of mental illness in other countries is similar to that for Australia, although rates may be higher in developed countries than in developing countries [16, 17].

1.2.1.2 Life course. Mental illness is commonly first experienced early in life, typically in late adolescence (18–24 years old) [20]. Two thirds of Australians who ever experience a mental disorder do so before the age of 21 [20]. The early development of mental illness is associated with subsequent detrimental social and economic impacts across the life course. Compared with Australians with other major health conditions, a person with a mental illness is less likely to be in the labour force [25]. Rates of mental illness among people who are homeless or in prison are substantially higher than in the general population, with up to 75% of homeless adults and around 40% of prisoners having a mental illness [25]. There is also an overrepresentation of people with severe mental illness within these groups [25].

1.2.1.3 Help-seeking. From a public health perspective, one of the most challenging aspects of mental illness is that the majority of people (65%) with a mental illness do not seek professional help for their condition [20, 26]. The reasons for this are complex and arise from an interaction of personal and interpersonal factors, and from structural factors in the mental healthcare system. Stigma towards mental illness is an important factor [27], with willingness to access mental health services diminished by a fear of embarrassment and/or discrimination. Poor mental health literacy and a preference for self-reliance have also been observed as barriers to help-seeking [27].

1.2.1.4 Services and national policy. Structurally, there are issues with the Australian mental healthcare system, many of which are yet to be addressed [28]. For example, there is evidence that inequity in the accessibility of mental healthcare pervades the system, particularly in non-urban areas and for minority groups [29-32]. It has been argued that the allocation of funding to services is often determined by administrative convenience rather than a person-centred needs analysis [33]. More funds are allocated by state governments to specialised mental healthcare services in hospitals (\$2.1 billion) for people at crisis point than to community-based services (\$1.9 billion) which might support people in daily life and prevent issues escalating to crisis point [34].

There is no single solution to remedy these issues, and in a fiscally tight political environment the prospects of the allocation of funding sufficient to rectify the situation are unlikely. One potential approach to addressing these challenges is to introduce a range of cost-effective and person-centred reforms [33]. In 2014, a report commissioned by the Australian Government (the National Review of Mental Health Programmes and Services) made 25 recommendations regarding structural reform to the mental health sector grouped under nine strategic directions [35]. Three of these strategic directions are of particular relevance to this thesis. First, *to improve access to services and support through innovative technologies*; second, *to build workforce and research capacity to support systems change*; and third, *to empower and support self-care and implement a new model of stepped care across Australia*. In particular, the report recommended improving supply, productivity and access to the mental health peer workforce. It also recommended providing easy access to self-help options that support communities to support themselves and each other, and to improve ease of navigation for stepping through the mental health system.

The study of MHISGs is relevant to all three of the above areas identified in the National Review of Mental Health Programs and Services. Online peer support is essentially provided by a volunteer workforce. This workforce may be positively influenced by policies which support it to act as scaffolding (access and support through transitions) for other services. Determining if and how well MHISGs are suited to performing this role depends on understanding more about how they currently operate.

Before considering the current literature on MHISGs, the next two sections will provide important background information on the broader field of e-mental health and its advantages, and on peer support.

1.2.2 E-mental health

E-mental health encompasses a range of technology-based services that provide treatment and support to consumers and carers. Broadly, they can be grouped into five categories of services delivered through Internet or mobile technology: 1) health promotion, wellness promotion and psycho-education; 2) prevention and early intervention; 3) crisis intervention and suicide prevention; 4) treatment; and 5) recovery and peer support [36]. E-mental health resources can be used wherever people have access to a computer, a telephone and/or the Internet. They are delivered via websites, smart-phone apps, sensor-based monitoring devices, computers and telephones.

E-mental health services circumvent the barriers associated with traditional mental health services in three key respects.

1) Location. There are significant geographical barriers to accessing face-to-face services among some sub-groups of the population. By contrast, e-mental health services can be accessed from any location with access to the Internet. In Australia, for example, 86% of households have access to the Internet, including more than 97% of households with children under 15 [37]. This is particularly relevant to people residing in rural and remote areas for whom access to mental health services is often limited and can involve substantial travel [38], although the prevalence of Internet access in rural and remote areas is 11 percentage points lower than in urban areas [37]. Culturally, people in rural areas value self-reliance, a factor which may increase the appeal of self-help services [38]. Access to face-to-face services can also be difficult for young people whose ability to reach these services may be mediated by

adults [39]. Online communities may also be particularly useful for people who are socially isolated [40] and those for whom there are physical barriers to travel.

2) *Cost*. Some e-mental health services have been shown to be more cost effective than the equivalent face-to-face services [41]. While the initial development of any mental health service requires substantial funding, the ongoing maintenance and delivery costs of e-mental services are small relative to the scale on which they can be delivered. For example, a cost-utility analysis of the e-mental health program *myCompass* found that the cost per quality-adjusted-life-year delivered by the program was one fifth that of treatment with antidepressants and one tenth that of recommended treatment by a psychologist [36]. Consequently, e-mental health programs can be delivered at low or no cost to the user. Longstanding freely available e-mental health programs such as *MoodGYM* have been used by over 1,000,000 people [36].

3) *Stigma*. The stigma surrounding mental illness comprises attitudes and beliefs that people with mental disorders are weak and that the conditions do not constitute real medical illnesses [42]. Those with a mental illness frequently internalise this public stigma, leading to a reluctance to seek formal help for their condition [43]. Those considering accessing face-to-face services may fail to do so due to fear of being stigmatised by others, including their health practitioners [43]. The help-seeking decisions of young people in particular are affected by perceived stigma [39]. A critical aspect of e-mental health services is that many can be accessed anonymously, a factor which reduces perceived stigma as a barrier to accessing help.

Additionally, e-mental health programs do not require a referral and can be accessed at any time of day that they are needed. Access to such services may reduce presentations at more costly face-to-face services that operate 24 hours a day, such as emergency departments. While e-mental health is not necessarily an adequate substitute for people who need access to specific face-to-face services, they may be helpful in providing support until other more appropriate services can be accessed.

The unique advantages of e-mental health services place them in a strong position to facilitate help-seeking among people who are not inclined to seek help from traditional mental health services [44]. They may also facilitate pathways between conventional services and assist those with a mental disorder to transition between and throughout different levels of stepped-care [36]. In the broader context of the mental healthcare system, the potential of e-mental

health has not been fully realised [44] primarily due to a failure of implementation [45]. Thus, although there is substantial evidence of the effectiveness of e-mental health services such as iCBT [46-50], including evidence that it is as effective as face-to-face delivered therapies [51, 52], such resources have not been systematically implemented [45] or integrated within the mainstream mental health system.

In contrast, as will be discussed below, little is known about online peer support. Such groups have been implemented widely, but there is an absence of foundational research concerning the nature and efficacy of these groups.

1.2.3 Peer support

In the mental health sector, peer support refers to supportive exchanges between individuals with a shared understanding of mental illness from their own lived experience [53]. These exchanges comprise social support, including emotional, informational, instrumental and appraisal support [54]. Peer support can be naturally occurring in any context or it can be intentionally facilitated. The degree of structure and formality in the latter can vary from informal self-help groups organised in the community to professional one-on-one relationships between a peer worker and a consumer in a mental health service. The setting may be face-to-face, over the telephone, or online.

Researchers have classified the interpersonal dynamics of peer support into three type: 1) mutual help; 2) consumer-run services; and 3) consumers as mental health providers [55, 56]. *Mutual help* comprises interactions between peers in a relatively symmetrical relationship in that they are engaged in peer-support for the purpose of both providing and receiving support. Internet support groups are typically regarded as comprising mutual help [4, 57]. *Consumer-run services* comprise an asymmetrical relationship whereby a volunteer or a peer-worker who is paid for their expertise as a peer, provides support to a consumer. Generally, the peer is further progressed in their recovery and provides a mentoring role to support the consumer. *Consumers as mental health providers* comprises support provided by a peer who is a health professional in a role that does not require them to be a consumer.

Formally, the involvement of peer support in mental healthcare dates back to the 18th century; however, it has only been since the 1990s [58, 59] that peer support has become popular and widely adopted [60]. The origin of peer support stems from the consumer movement of the

1970s which advocated for mental health consumer rights and de-institutionalisation of mental healthcare [61]. The movement challenged the assumptions and negative consequences of traditional mental healthcare and brought peers together with the aim of creating their own empowering alternatives to existing care [62]. It has been suggested that peer support and peer-run services provide three unique contributions not offered by traditional (non-consumer) mental healthcare [60, 63]. These are:

- 1) *instilling hope* through positive self-disclosure: peers can empower others by demonstrating that recovery is possible and encourage them to strive towards it;
- 2) peers providing a *role model* for others to demonstrate how to navigate day-to-day life issues, including for example dealing with social, financial and healthcare system related issues or discrimination [64]; this is largely achieved by imparting experiential knowledge; and
- 3) forming a *valued relationship* that is characterised by trust, respect, acceptance, empathy, conditional regard and encouragement; such a relationship is seen as critical to engaging people to assist and motivate them to pursue their recovery [65, 66].

An important aspect of the consumer movement which influenced the development of peer support services was an alternative model of recovery that was used to define health and wellbeing [59]. Prior to the emergence of the consumer movement, the prevailing model of recovery was the biomedical model [67]. This model framed recovery objectively and conceptualised health as a state to which a person could be returned after the symptoms of illness had been eliminated (clinical recovery) [68, 69]. Since the rise of the consumer movement and with the increasing use of peer workers in mental healthcare, a second model has gained traction [59]. In contrast to the medical model, in which recovery is framed as 'recovery from' mental illness, the consumer model frames recovery as 'recovery in' mental illness [70]. In the consumer model, a greater emphasis is placed on helping a person to reach meaningful self-defined goals, as well as supporting hope, empowerment and social connection (personal recovery) [69, 71]. This model has attracted criticism from some clinicians for lacking a theory-driven approach [72]. Measures of personal recovery developed with consumer input, such as the Recovery Assessment Scale [73], have been found to correlate with self-esteem, self-empowerment, quality of life, symptoms, and positive relationships with social support [74]; but generally not with traditional clinical measures or recovery or functioning [75] such as the Health of the Nation Outcome Scales [76].

While the origins of peer support are strongly associated with the consumer model of recovery, it is not necessary for a participant to adopt the recovery model in order to participate in a peer-support group, particularly in online support groups that are freely and publically accessible. A person may adhere to a belief in either the consumer or the clinical model of recovery, or neither [77]. This distinction between the two models is of more than theoretical interest. The findings from the current thesis [78] suggest that the historical [61] and ongoing [72] tension between the two models may be reflected in the social dynamics of a mental health peer-support group which in turn may affect the nature of participation in these groups.

The evidence regarding the effectiveness of face-to-face peer support is promising, but mixed and not of high quality. A systematic review of randomised controlled trials involving peer support in community mental health services observed different findings across the three types of peer support [56]. For mutual help, low-grade evidence suggested there are positive effects on quality of life and empowerment, and a reduction in symptoms of depression and anxiety [56]. For consumer-run services, low-grade evidence suggested that there are positive effects for self-reported recovery and hope [56]. For consumers as mental health providers (as compared with standard care), no evidence of positive effects was observed [56]. Five previous systematic reviews have found more positive results [79-83]; however, these reviews all included low-quality trials. In general, the findings of studies on peer support are regarded as promising, but require further high-quality research.

1.2.4 Peer support on the Internet

The prevalence of participation in online health-related peer support has become increasingly common over recent decades. A survey of 9,187 people in the USA, conducted in 2003, found that the lifetime prevalence of participation in an Internet support group was 1.5%. A survey of 3,001 people in the USA in 2011 found that 18% of Internet users overall and 23% of Internet users with a chronic health condition had searched online in the last year to find other people with similar health issues [1]. A survey in 2013 found that 8% of all Internet users had actively engaged in peer-support groups by either posting a question or sharing information based on their personal health experience [2]. A 2016 survey of 2,670 people found that 78% of newly diagnosed patients reported searching for information or support from people online [84]. The 2003 survey found that mental illness was the most frequent health condition for which people were seeking online peer support, ahead of stroke, diabetes, cancer and arthritis

[85]. It has been suggested that the higher usage of online peer support for mental illness may be attributable to the stigma associated with mental illness [86].

With the emergence of social media as a prominent communication platform, online health-related peer support now exists in many different formats and is characterised by substantial heterogeneity [87]. In addition to the many stand-alone Internet support groups, increasingly large numbers of supportive communities and person-centred networks have developed organically in social media channels, such as in *Facebook* groups [88], around *Twitter* hashtags [89] and in the comment threads of *YouTube* channels [90] and blogs [91]. In addition, entire social network platforms have been designed for consumers to record their daily illness and treatment experiences, e.g. *Patients Like Me* [92]. Some of these communities are moderated and managed by professionals, consumers or a combination of the two, but many are not moderated. A consensus on the nomenclature for these groups has not yet been clearly established. Generally, Online Health Community (OHC) is used as the most broadly encompassing term to describe any digital social network sharing health-related information and the term Internet Support Group (ISG) is used to describe a website, such as a forum, bulletin board, chatroom or newsgroup, dedicated specifically to peer support for a certain health condition.

Most research to date has focused on ISGs. Despite differences in structure and functionality, fundamentally ISGs and OHCs comprise online social networks. It is therefore possible that evidence about ISGs can inform the broader field of OHCs and vice versa. However, at this stage, it is not clear. Therefore the studies in the current thesis focus on systematically investigating participation in one format, an ISG, for which the applicability of the findings to other formats may be investigated by future research. ISGs present a convenient means to analyse interpersonal interactions because, in contrast with face-to-face groups, the data is automatically collected by the operating software of an ISG. This allows for the use of innovative methods of analysis that are systematic, scalable and replicable.

1.2.5 Mental health Internet support groups

There are two main streams of research on Mental Health Internet Support Groups (MHISGs). One research stream is focused on investigating the efficacy of MHISGs [10] and the other on the nature of MHISGs such as the characteristics of users and the content of their interactions [4].

1.2.5.1 Efficacy. Two systematic reviews of ISG efficacy, both of which reported on depressive symptom reduction, concluded that there was not, at the time of the reviews, sufficient evidence to indicate positive effects [9, 10]. A systematic review of MHISGs for young people reached a similar conclusion [93]. Some researchers question the validity of randomised controlled trials (RCTs) to investigate the efficacy of ISGs, asserting that the interventions in such studies lack the therapeutic components of a 'real' ISG [94]. However, since the publication of the two systematic reviews of adult ISGs, an RCT involving an MHISG that was established for research has found evidence of its efficacy for symptom reduction [95, 96]. The MHISG in this trial was modelled on the MHISG used in this thesis (BlueBoard), but throughout the study it was not visible or accessible by BlueBoard users. It was closed to all but the study participants. Moreover, unlike BlueBoard, research participants were given instructions on how to participate including engaging in mutual support by writing four posts per week on different topics each week for 12 weeks. This included topics such as 'What helps? Who helps?', 'General chit chat', and views on anti-depressants and psychological therapies [95]. Relative to participants in an attention-control condition comprising a health-information website, no significant difference was found at post-test. However, at 6- and 12-month follow-ups, MHISG participants were found to have significantly lower depressive symptom severity relative to control group participants.

Recently, a randomised trial evaluated the impact of directing people who self-reported mental illness to participate in a fully-mature MHISG (Psych Central) compared with direction to participate in an expressive writing activity. There was no evidence that the MHISG was more effective than the expressive writing in reducing depressive and anxiety symptoms. Reductions in both groups were similar to those observed in no-intervention control groups reported in a meta-analysis of computer-based mental health treatments. In attempting to interpret this finding in the context of previous positive trials and the broader literature on MHISGs, including their growing popularity, the authors referred to this finding as 'paradoxical' and concluded that directing people to participate in MHISGs could not be recommended [97].

Beyond clinical efficacy, it is thought that MHISGs, like face-to-face peer support groups, have the potential to yield outcomes that align with the consumer model of recovery, for example by empowering participants to take a more active role in their recovery by fostering a sense of hope, self-control, confidence and independence [98]. It is therefore common for ISGs to be provided in conjunction with other psychoeducational and therapeutic Internet interventions with the aim of promoting engagement [99]. The aforementioned trial involving a depression

ISG modelled on BlueBoard found that participants of the depression ISG had a significantly higher sense of empowerment at post-test and self-esteem at a 12-month follow-up relative to participants in the control condition [100]. Other studies have surveyed participants and compared the responses of people who actively participate (i.e., post) versus those who 'lurk' (i.e., only read posts). It has been found that active participants report more benefits of participation than people who lurk, including emotional support, helping other consumers, expressing their emotions [101] and greater social wellbeing [102]. However, such differences might reflect pre-existing differences between active and passive participants rather than a differential effect of passive versus active exposure.

1.2.5.2 Nature. A broad range of research studies have documented the nature of MHISGS. There has been one systematic review of studies concerning the nature of depression ISGs [4]. It reported that most studies in this field have been descriptive and predominantly focused on describing the characteristics of users, the types of content comprising their interactions and site-usage statistics. The review concluded that research in this area has typically relied upon data sourced from surveys or manual content analyses highlighting the potential bias of such findings given selection effects associated with the methodology. The authors recommended that studies use data collected at the point of registration in order to provide higher-quality findings regarding user characteristics.

The above review found that users were most commonly consumers who were currently experiencing depression, aged in their mid-20s to 40s and that there was mixed evidence for gender composition [4]. Two studies in this review reported that the majority of survey respondents scored above the cut-off for depression on the CES-D (1–2 months after joining the group) [103] or had current major depression as measured by the Major Depression Inventory [104] and reported that they had received a diagnosis of depression [103]. Use of treatments and services was found to vary between 26% of depressed users [104] to 92% of all users [103]; 36% reported that participation in the ISG had been a catalyst for formal help-seeking [104]. More recently, a survey of users across a range of MHISGs has produced similar findings [105]

Studies of users' post content have predominantly used pre-formulated coding schemes to identify the prevalence of various types of social support or other types of content such as requests for help [4]. The authors of the systematic review observed that differences between the coding schemes used in different studies made it difficult to synthesise results across

publications but concluded that it was clear that posts comprising emotional support (mean across studies = 34%) and information support (mean = 26.2%) were common [4]. Findings from a singular study identified that experiential knowledge (14%) and self-disclosure (50.6%) were also common, but that only a minority of posts contained requests for help (13%) and second-hand professional knowledge (3%) [57]. One other study reported that almost a quarter of posts involved social companionship (e.g. chit-chat) content [106]. There is evidence that relative to ISGs for other health conditions, depression ISGs contain more content comprising emotional support, less informational support and less content about medical treatments and procedures [107]. More recently, an analysis of BlueBoard post content following an inductive approach investigated user-perceived advantages and disadvantages of participation [108]. Positive personal change in terms of emotional, cognitive and behavioural effects were the most common (74.0% of statements), followed by valued social interactions (65.3% of statements). Content included all posts in the MHISG sampled consecutively for three months and is therefore likely to have over-represented the views of a minority of high-contributing users.

1.2.6 Active participation in mental health Internet support groups

Active participation, as opposed to passive participation, is any activity in the group that is visible to others. Passive participation involves viewing without contributing. In accordance with the power law that describes the distribution of engagement, the number of people participating passively far exceeds the number of people participating actively. The current research is focused solely on active participants.

A systematic review conducted at the outset of this thesis (Chapter 2) analysed studies on the nature of active participation (participation styles) in an MHISG. Since very few studies have been conducted on MHISGs, this review of participation styles was expanded to include all online health communities. There are three aspects of any user's active participation in any online health community: *activity*, *network connectivity* and *content*. *Activity* refers to the type of act performed by a user (e.g. writing a post, starting a thread or commenting on another user's post). *Network connectivity* refers to whom the activity was directed or with whom it was shared. *Content* refers to the type of content that was included in the activity, that is, what was written. The majority of studies on MHISGs used one-dimensional measures of participation, particularly activity-based metrics, as measures of engagement. None of the studies identified by the systematic review proposed an overall model for the nature of

participation, other than those which reported that the distribution of engagement could be grouped by a rule of thumb known as the 90-9-1 principle [13], illustrated in figure 1.1, in which 90% of users are labelled 'lurkers', 9% 'contributors' and 1% 'superusers'. In fact, as demonstrated in the study reported in Chapter 3, this activity is better characterised by a power law, referred to as Zipf's law [12], and a better fit of the distribution to Zipf's law may reflect a greater cohesiveness of the community [12].

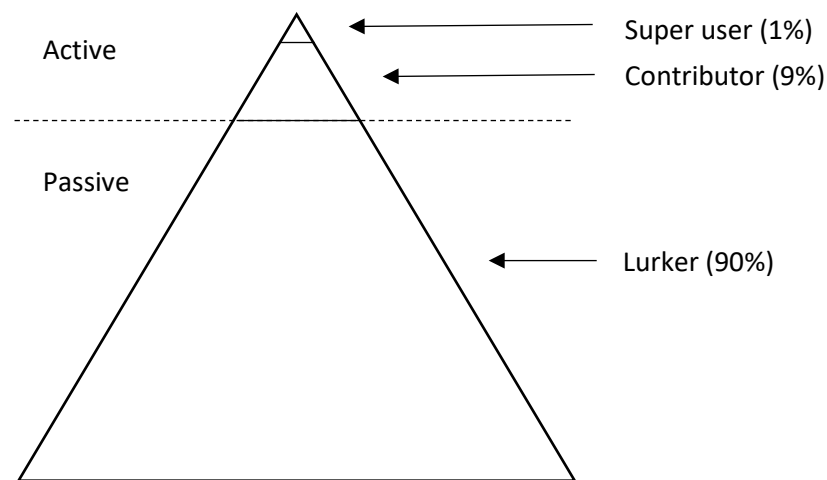


Figure 1.1. The 90-9-1 principle of active participation in an online community reproduced from Nielsen et al. [15].

There is, however, one model of active participation in an MHISG [109] that was not included in the Chapter 2 systematic review as it did not fulfil the study inclusion criteria, being based on a cross-sectional self-report survey rather than on data from an empirical study of the participation itself. The model, referred to as 'the membership life-cycle' [109] and illustrated in figure 1.2, describes universal stages through which users traverse during their membership. It describes all users as 'distressed newcomers' who become either 'active help receivers' or 'passive followers', followed by 'relieved survivors' who either leave as 'exiting quitters' or go on to become 'active help providers' for other users. The cross-sectional survey that was the basis for this model included respondents from 16 different depression ISGs of whom 40% had been members for over a year [109]. The study found significant correlations between longer membership duration and stronger endorsement of statements about benefits of participation: 'I cope with the depression better', 'I am more capable in dealing with daily

tasks' and 'I can be of help to others'. The credibility of this model is limited by the potential for selection bias in the survey and the lack of direct, longitudinal observation of user participation. It is underpinned by an assumption that MHISGs may be effective only through a single mechanism of action, although there is no evidence for this. Thus, there is a clear need to undertake research to generate empirical data to investigate the validity of such assumptions and to generate an accurate model of MHISGs.

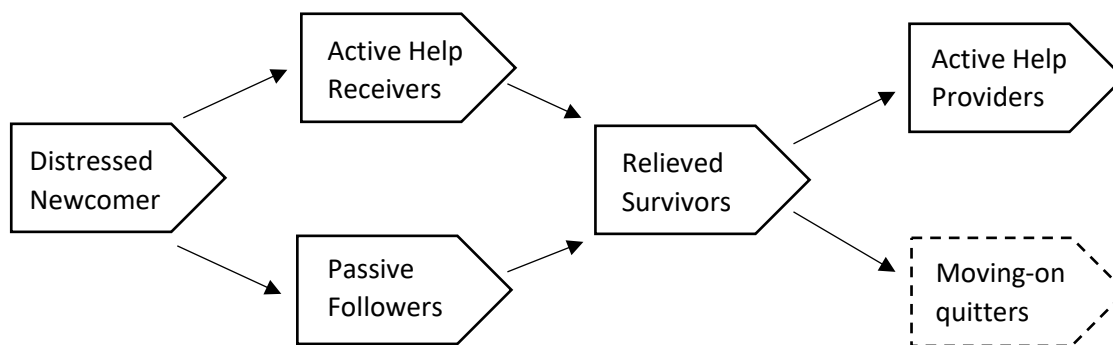


Figure 1.2. The membership life-cycle model in a depression Internet support group reproduced from Nimrod et al. [109].

1.3 The current research

1.3.1 Aims and overview of the studies in the thesis

The current research comprises a series of exploratory studies which aimed to address fundamental gaps in extant knowledge regarding participation in MHISGs. The work aimed to both document in detail the nature of active participation by the MHISG user community, and to extend the current research on engagement to investigate participation inequality and its association with user characteristics, support group network connectivity and post content. The specific questions addressed by the work in the thesis are summarised in box 1.1 below.

The systematic review (Chapter 2 [11]) sought to document the current state of knowledge about the elements of active participation, with the exception of user characteristics (e.g.,

demographic status), a factor that has been addressed by a previous systematic review [4]. Thereafter, each of the empirical studies included in the thesis (Chapters 3-7) addressed one or more aspects of participation. With the exception of Chapter 3, the empirical studies are similar in structure. Firstly, each study investigates the target aspect of participation for all active users combined. Each study then examines whether users differ as a function of their level of engagement. As the overall investigation unfolded, the designs of consecutive studies were informed by findings from previous studies. Linking statements in the form of forewords to each chapter explain this development. To avoid duplication, the paragraph which follows is restricted to a descriptive overview of the content of each chapter and study.

The first empirical study [12], presented in Chapter 3, investigated user activity and provides a context for the subsequent empirical work. The aim of this study was to clarify and mathematically refine the description of the distribution of engagement that had been previously assumed to follow a broadly specified rule of thumb. Chapter 4 presents a study involving an analysis of BlueBoard usage and users characteristics [110]. This included an analysis of the association between user characteristics and user activity to determine whether higher-engaged users differed from lower-engaged users in their demographic characteristics. Chapter 5 presents the findings from of an analysis of the network connectivity of users using a modularity algorithm to determine sub-groupings of users in the BlueBoard community structure [111]. A further analysis was undertaken to determine whether membership in the modules (sub-groups) was associated with commonalities in the user's characteristics, and whether higher usage was systematically associated with the characteristics that defined these commonalities. In Chapter 6, the content of posts across the entire forum was modelled with a machine-learning algorithm to determine if there were systematic differences between the content of posts written by higher- compared to lower-engaged users, and whether this changed when higher-engaged users were communicating with other users as opposed to between themselves [78]. In Chapter 7, user characteristics were revisited using a qualitative analysis to investigate the content of user's initial posts to determine if there were systematic differences apparent in the way that users presented themselves.

Box 1.1: Research questions

RQ 1: In what ways has participation previously been defined and measured in an MHISG, and other online health communities? (Chapter 2)

RQ 2: What is the mathematical distribution of engagement as measured by posting frequency among active users in an MHISG? (Chapter 3)

RQ 3

a: What are the characteristics of BlueBoard users? (Chapter 4)

b: Are there systematic differences in the characteristics of users with different levels of engagement? (Chapter 4)

RQ 4

a: Are there sub-groups (modules) within the BlueBoard community structure that comprise users who communicate with each other relatively more frequently than with other users? (Chapter 5)

b: Do users within the same sub-group share a higher degree of common characteristics, including posting frequency in disorder-specific forums, than would be expected by chance? (Chapter 5)

c: Within each sub-group and with respect to the previously ascertained common characteristic, are there systematic differences in that characteristic between higher- and lower-engaged users? (Chapter 5)

RQ 5

a: About what topics do BlueBoard users communicate? (Chapter 6)

b: Are there systematic differences in the degree to which users with different levels of engagement write content pertaining to each of these topics? (Chapter 6)

c: Do these differences change when higher-engaged users are communicating with lower-engaged users as opposed to communicating within their own grouping? (Chapter 6)

RQ 6

a: How do BlueBoard users present themselves in their initial post? (Chapter 7)

b: Are there systematic differences in the way that users with different levels of engagement present themselves? (Chapter 8)

Together, these studies aimed to provide a basis for the development of an evidence-based model of active participation in an MHISG. Through its focus on systematic differences between higher- and lower-engaged users, this research also sought to provide evidence with potential implications for community management, the design of future research and the formulation of policy regarding MHISGs. For MHISG management, the novel methods of analysis, such as the network modularity algorithm and topic model algorithm, aimed to provide a perspective enabling visualisation of large-scale participatory patterns that may not otherwise be apparent to the community managers. Further, this research aimed to provide information regarding systematic differences between users, particularly those which are apparent from the outset of their participation that may assist in the tailoring of service delivery strategies that are optimised for individual differences between users. With respect to future research, the current research sought to provide evidence that might inform the design of studies seeking to measure the effects of participation to ensure analyses take into account any differences in patterns of activity between underlying sub-groups which if ignored could lead to misleading conclusions. Finally, with respect to policy, this research had the potential to provide a new perspective on the role that is performed by MHISGs in the mental health sector and highlight opportunities for service delivery innovation. In Chapter 8, the thesis concludes by presenting a model of active participation and a discussion of implications in above three domains of practice, research and policy.

1.3.2 BlueBoard

The current research, with the exception of the systematic review in Chapter 2, has been conducted using data from the MHISG BlueBoard. BlueBoard was established by the Centre for Mental Health Research at the Australian National University to operate as an MHISG with informed consent built into the registration process and human research ethics clearance so that all data collected from the activities of participants could be used for research purposes. BlueBoard was first established in 2003 as a mood disorder group. It was closed in 2007 and 2008 due to lack of funding and re-established on 1 October 2008, but closed again on 30 June 2016 with the retirement of its founder. Data for the current research included all posts generated between 1 October 2008 and 23 May 2014 (131,004 posts made by 2,934 users) and the characteristics of users recorded during registration (age, gender, location and status as either consumer, carer or other).

BlueBoard was moderated by consumers who were trained and supervised by a registered psychologist. Moderators did not actively participate in any of the forums. Rather they enforced rules (e.g., by editing posts to remove any personally identifying information or inappropriate content such as discussion of suicide methods) in order to manage risk and maintain a safe environment for users. Moderators alerted the responsible user via a private notification. General notices and other informational posts (352 posts made by 10 moderators) were removed from all analyses.

BlueBoard comprised 10 forums in which users communicated about a range of mental health issues, including: eight condition-specific forums (Depression, Bipolar Disorder, Generalised Anxiety, Social Anxiety, Borderline Personality Disorder, Eating Disorder, Panic Disorder and Obsessive Compulsive Disorder); one forum dedicated to carers of people with mental health issues; and one forum for general discussion. All forums were created on 1 October 2008 with the exception of the Obsessive Compulsive Disorder, Borderline Personality Disorder and Eating Disorder forums which were added on 1 June 2009, 1 March 2010 and 30 July 2012 respectively. These additional forums were created in response to requests from users.

BlueBoard users comprised members of the general public. Advertising of BlueBoard occurred mainly via links from online mental health information hubs such as bluepages.anu.edu.au and other depression websites. Following the establishment of BlueBoard, postcard flyers were mailed to general practitioners for display in waiting rooms. Postcards were also distributed at professional conferences. BlueBoard was delivered on the vBulletin platform which included features similar to other ISGs such as the ability to quote other users in posts and the provision of user information (total posts and registration date) displayed beside the username of each post. It did not include labels for users recognising different levels of contributions to the forum. In addition, to preclude users exchanging contact details and posting prohibited and potentially inappropriate links and immediately deleting them, the editing facility was not available to users and the script underlying the platform was modified to prevent users from inserting hyperlinks.

1.4 References

1. Fox S. Peer-to-peer Health Care. Washington, D.C.: PewResearchCentre, 2011.
2. Fox S, Duggan M. Health Online 2013. Washington, D.C.: PewResearchCentre, 2013.
3. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of Medical Internet Research*. 2013; 15(4): e85.
4. Griffiths KM, Calear AL, Banfield MA, Tam A. Systematic review on Internet Support Groups (ISGs) and depression (2): What is known about depression ISGs? *Journal of Medical Internet Research*. 2009; 11(3): e41.
5. Fox S. After Dr Google: peer-to-peer health care. *Pediatrics*. 2013; 131(Supplement 4): S224-S5.
6. Griffiths F, Cave J, Boardman F, Ren J, Pawlikowska T, Ball R, et al. Social networks—the future for health care delivery. *Social Science & Medicine*. 2012; 75(12): 2233-41.
7. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*. 2013; 35(4): 332-8.
8. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences*. 2016; 25(02): 113-22.
9. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *British Medical Journal*. 2004; 328(7449): 1166.
10. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? *Journal of Medical Internet Research*. 2009; 11(3): e40.
11. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From help-seekers to influential users: a systematic review of participation styles in online health communities. *Journal of Medical Internet Research*. 2015; 17(12): e271.
12. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions*. 2014; 1(4): 165-8.
13. van Mierlo T. The 1% rule in four digital health social networks: an observational study. *Journal of Medical Internet Research*. 2014; 16(2): e33.
14. van Mierlo T, Hyatt D, Ching AT. Mapping power law distributions in digital health social networks: methods, interpretations, and practical implications. *Journal of Medical Internet Research*. 2015; 17(6): e160.
15. Nielsen J. Participation Inequality: Lurkers vs Contributors in Internet Communities 2014 [05/06/2014]. Archived at <http://www.webcitation.org/6Q7EwEncA>. Available from: <http://www.nngroup.com/articles/participation-inequality/>.
16. Kessler RC, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J, et al. The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiologia e Psichiatria Sociale*. 2009; 18(1): 23-33.
17. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980–2013. *International Journal of Epidemiology*. 2014; 43(2): 476-93.
18. World Health Organization. Mental Health Atlas. Geneva: World Health Organization, 2011.
19. World Health Organization. Comprehensive Mental Health Action Plan 2013–2020. Geneva: World Health Organization, 2013.

20. Australian Bureau of Statistics. National survey of mental health and wellbeing: Summary of results 2007. cat. no. 4326.0. Canberra: Australian Bureau of Statistics, 2008.
21. Australian Institute of Health and Welfare. Australian Burden of Disease Study: Impact and causes of illness and deaths in Australia 2011. Canberra: Australian Institute of Health and Welfare, 2016.
22. The Royal Australian and New Zealand College of Psychiatrists. The economic cost of serious mental illness and comorbidities in Australia and New Zealand. Melbourne: The Royal Australian and New Zealand College of Psychiatrists, 2016.
23. Australian Bureau of Statistics. Australian Social Trends – Mental health cat. no. 4102.0. . Canberra: Australian Bureau of Statistics, 2009.
24. Australian Government Department of Health. The Mental Health of Australians 2: 2007 National Survey of Mental Health and Wellbeing highlights. Canberra: Australian Government Department of Health, 2009.
25. Australian Government Department of Health. Fourth national mental health plan: an agenda for collaborative government action in mental health 2009-2014. Canberra: Australian Government Department of Health, 2009.
26. Australian Government Department of Health and Ageing. National Mental Health Report 2013. Canberra: Australian Government Department of Health and Ageing, 2013.
27. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*. 2010; 10(1): 113.
28. Griffiths KM, Mendoza J, Carron-Arthur B. Whereto mental health reform in Australia: is anyone listening to our independent auditors. *Medical Journal of Australia*. 2015; 202(4): 172-4.
29. Aisbett DL, Boyd CP, Francis KJ, Newnham K, Newnham K. Understanding barriers to mental health service utilization for adolescents in rural Australia. *Rural and Remote Health*. 2007; 7(624): 1-10.
30. Brown A, Rice SM, Rickwood DJ, Parker AG. Systematic review of barriers and facilitators to accessing and engaging with mental health care among at - risk young people. *Asia - Pacific Psychiatry*. 2016; 8(1): 3-22.
31. Moore T, Sutton K, Maybery D. Rural mental health workforce difficulties: a management perspective. *Rural and Remote Health*. 2010; 10(3): 1519.
32. Vasey K, Manderson L. Regionalizing immigration, health and inequality: Iraqi refugees in Australia. *Administrative Sciences*. 2012; 2(1): 47-62.
33. Mendoza J, Elson A, Gilbert Y, Bresnan A, Rosenberg S, Long P, et al. Obsessive Hope Disorder: Reflections on 30 years of mental health reform in Australia and visions for the future. Caloundra, QLD: Connetica, 2013.
34. Australian Institute of Health and Welfare. Mental Health Services in Australia. Canberra: Australian Institute of Health and Welfare, 2016.
35. National Mental Health Commission. Contributing Lives, Thriving Communities - Review of Mental Health Programmes and Services. Canberra: National Mental Health Commission, 2014.
36. Christensen H, Proudfoot J, Woodward A, Hosie A, Klein B, Morgan C, et al. e-Mental health services in Australia 2014: Current and future. Canberra: The e-Mental Health Alliance, 2014.
37. Australian Bureau of Statistics. Household Use of Information Technology, Australia, 2014-15. cat no. 8146.0. Canberra: Australian Bureau of Statistics, 2016.
38. Griffiths KM, Christensen H. Internet - based mental health programs: A powerful tool in the rural medical kit. *Australian Journal of Rural Health*. 2007; 15(2): 81-7.
39. Cohen A, Medlow S, Kelk N, Hickie I. Young people's experiences of mental health care: Implications for the headspace National Youth Mental Health Foundation. *Youth Studies Australia*. 2009; 28(1): 13.

40. Horrigan J. Online communities: Networks that nurture long-distance relationships and local ties. Washington, D.C.: Pew Internet & American Life Project, 2001.
41. Hedman E, Ljótsson B, Lindefors N. Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2012; 12(6): 745-64.
42. Griffiths KM, Christensen H, Jorm AF. Predictors of depression stigma. *BMC Psychiatry*. 2008; 8(1): 25.
43. Barney LJ, Griffiths KM, Jorm AF, Christensen H. Stigma about depression and its impact on help-seeking intentions. *Australian and New Zealand Journal of Psychiatry*. 2006; 40(1): 51-4.
44. Christensen H, Hickie IB. Using e-health applications to deliver new mental health services. *Medical Journal of Australia*. 2010; 192(11): S53.
45. Meurk C, Leung J, Hall W, Head BW, Whiteford H. Establishing and governing e-mental health care in Australia: a systematic review of challenges and a call for policy-focussed research. *Journal of Medical Internet Research*. 2016; 18(1): e10.
46. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cognitive Behaviour Therapy*. 2009; 38(4): 196-205.
47. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PloS One*. 2010; 5(10): e13196.
48. Cuijpers P, Marks IM, van Straten A, Cavanagh K, Gega L, Andersson G. Computer - aided psychotherapy for anxiety disorders: a meta - analytic review. *Cognitive Behaviour Therapy*. 2009; 38(2): 66-82.
49. Griffiths KM, Farrer L, Christensen H. The efficacy of Internet interventions for depression and anxiety disorders: a review of randomised controlled trials. *Medical Journal of Australia*. 2010; 192(11): S4-S11.
50. Spek V, Cuijpers PIM, Nyklíček I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychological Medicine*. 2007; 37(03): 319-28.
51. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet - based vs. face - to - face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta - analysis. *World Psychiatry*. 2014; 13(3): 288-95.
52. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. *PLoS One*. 2014; 9(7): e100674.
53. Repper J, Carter T. Using personal experience to support others with similar difficulties. London: Together, 2010.
54. Heaney CA, Israel BA. Social networks and social support. *Health Behavior and Health Education: Theory, Research, and Practice*. 2008; 4: 189-210.
55. Davidson L, Chinman M, Kloos B, Weingarten R, Stayner D, Tebes JK. Peer support among individuals with severe mental illness: A review of the evidence. *Clinical Psychology: Science and Practice*. 1999; 6(2): 165-87.
56. Lloyd-Evans B, Mayo-Wilson E, Harrison B, Istead H, Brown E, Pilling S, et al. A systematic review and meta-analysis of randomised controlled trials of peer support for people with severe mental illness. *BMC Psychiatry*. 2014; 14(1): 39.
57. Salem DA, Bogat GA, Reid C. Mutual help goes on-line. *Journal of Community Psychology*. 1997; 25(2): 189-207.
58. Anthony WA. Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal*. 1993; 16(4): 11.

59. Frese FJ, Davis WW. The consumer–survivor movement, recovery, and consumer professionals. *Professional Psychology: Research and Practice*. 1997; 28(3): 243.
60. Davidson L, Bellamy C, Guy K, Miller R. Peer support among persons with severe mental illnesses: a review of evidence and experience. *World Psychiatry*. 2012; 11(2): 123-8.
61. Chamberlin J. The ex-patients' movement: Where we've been and where we're going. *The Journal of Mind and Behavior*. 1990; 11(3): 323-36.
62. McLean A. Empowerment and the psychiatric consumer/ex-patient movement in the United States: Contradictions, crisis and change. *Social Science & Medicine*. 1995; 40(8): 1053-71.
63. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. *Psychiatric Rehabilitation Journal*. 2004; 27(4): 392.
64. Verhaeghe M, Bracke P, Bruynooghe K. Stigmatization and self-esteem of persons in recovery from mental illness: the role of peer support. *International Journal of Social Psychiatry*. 2008; 54(3): 206-18.
65. Chinman M, Young AS, Hassell J, Davidson L. Toward the implementation of mental health consumer provider services. *The Journal of Behavioral Health Services & Research*. 2006; 33(2): 176.
66. Mowbray CT, Moxley DP, Thrasher S, Bybee D, Harris S. Consumers as community support providers: Issues created by role innovation. *Community Mental Health Journal*. 1996; 32(1): 47-67.
67. Engel GL. The need for a new medical model: a challenge for biomedicine. *Holistic Medicine*. 1989; 4(1): 37-53.
68. Peyser H. Commentary: What is recovery? A commentary. *Psychiatric Services*. 2001; 52(4): 486-7.
69. Slade M, Amering M, Oades L. Recovery: an international perspective. *Epidemiologia e Psichiatria Sociale*. 2008; 17(02): 128-37.
70. Davidson L, Roe D. Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health*. 2007; 16(4): 459-70.
71. Jacobson N, Greenley D. What is recovery? A conceptual model and explication. *Psychiatric Services*. 2001; 52(4): 482-5.
72. Davidson L, O'Connell M, Tondora J, Styron T, Kangas K. The top ten concerns about recovery encountered in mental health system transformation. *Psychiatric Services*. 2006; 57(5): 640-5.
73. Corrigan PW, Salzer M, Ralph RO, Sangster Y, Keck L. Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin*. 2004; 30(4): 1035.
74. Corrigan PW, Gifford D, Rashid F, Leary M, Okeke I. Recovery as a psychological construct. *Community Mental Health Journal*. 1999; 35(3): 231-9.
75. Andresen R, Caputi P, Oades LG. Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research*. 2010; 177(3): 309-17.
76. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. *The British Journal of Psychiatry*. 1998; 172(1): 11-8.
77. Bird V, Leamy M, Tew J, Le Boutillier C, Williams J, Slade M. Fit for purpose? Validation of a conceptual framework for personal recovery with current mental health consumers. *Australian and New Zealand Journal of Psychiatry*. 2014; 48(7): 644-53.
78. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Griffiths KM. What's all the talk about? Topic modelling in a mental health Internet support group. *BMC Psychiatry*. 2016; 16(1): 367.
79. Pfeiffer PN, Heisler M, Piette JD, Rogers MAM, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. *General hospital psychiatry*. 2011; 33(1): 29-36.

80. Pistrang N, Barker C, Humphreys K. Mutual help groups for mental health problems: A review of effectiveness studies. *American Journal of Community Psychology*. 2008; 42(1-2): 110-21.
81. Pitt V, Lowe D, Hill S, Prictor M, Hetrick SE, Ryan R, et al. Consumer-providers of care for adult clients of statutory mental health services. *Cochrane Database of Systematic Reviews*. 2013; 3(9): Article Number CD004807.
82. Repper J, Carter T. A review of the literature on peer support in mental health services. *Journal of Mental Health*. 2011; 20(4): 392-411.
83. Simpson EL, House AO. Involving users in the delivery and evaluation of mental health services: systematic review. *British Medical Journal*. 2002; 325(7375): 1265.
84. Janssen Pharmaceuticals Inc. Janssen Announces HealtheVoices™ Impact Fund: PRNewswire; 2017 [16/06/2017]. Archived at <http://www.webcitation.org/6rFiTgFOy>. Available from: <http://www.prnewswire.com/news-releases/janssen-announces-healthevoices-impact-fund-300444015.html>.
85. Owen JE, Boxley L, Goldstein MS, Lee JH, Breen N, Rowland JH. Use of health - related online support groups: population data from the California health interview survey complementary and alternative medicine study. *Journal of Computer - Mediated Communication*. 2010; 15(3): 427-46.
86. Davison KP, Pennebaker JW, Dickerson SS. Who talks? The social psychology of illness support groups. *American Psychologist*. 2000; 55(2): 205.
87. Orizio G, Schulz P, Gasparotti C, Caimi L, Gelatti U. The world of e-patients: A content analysis of online social networks focusing on diseases. *Telemedicine and e-Health*. 2010; 16(10): 1060-6.
88. Hartland T, Cummings J. Coming Together to Prevent Youth Suicide 2017 [28/01/2017]. Available from: <https://www.facebook.com/groups/preventingyouthsuicide/>
89. Young C. Community management that works: how to build and sustain a thriving online health community. *Journal of Medical Internet Research*. 2013; 15(6): e119.
90. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. *PLoS One*. 2014; 9(10): e110171.
91. Counselling A. The 34 best bloggers who advocate for mental health and wellness 2017 [16/06/2017]. Archived at <http://www.webcitation.org/6rFj4DHDv>. Available from: <https://www.australiacounselling.com.au/mental-health-articles/best-bloggers-mental-health/>.
92. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research*. 2010; 12(2): e19.
93. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. *JMIR Mental Health*. 2015; 2(2): e19.
94. Barak A, Grohol JM, Pector E. Methodology, validity, and applicability: A critique on Eysenbach et al. *British Medical Journal*. 2004; 328: 1166.
95. Griffiths KM, Crisp D, Christensen H, Mackinnon AJ, Bennett K. The ANU WellBeing study: a protocol for a quasi-factorial randomised controlled trial of the effectiveness of an Internet support group and an automated Internet intervention for depression. *BMC Psychiatry*. 2010; 10: 20.
96. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. *PLoS One*. 2012; 7(12): e53244.
97. Dean J, Potts HWW, Barker C. Direction to an Internet support group compared with online expressive writing for people with depression and anxiety: a randomized trial. *JMIR Mental Health*. 2016; 3(2): e12.

98. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. *Computers in Human Behavior*. 2008; 24(5): 1867-83.
99. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to Internet-delivered healthy lifestyle promotion interventions? A systematic review. *Journal of Medical Internet Research*. 2011; 13(1): e2.
100. Crisp D, Griffiths K, Mackinnon A, Bennett K, Christensen H. An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. *Psychiatry Research*. 2014; 216(1): 60-6.
101. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. *Journal of Medical Internet Research*. 2011; 13(4): e122.
102. van Uden-Kraan C, Drossaert C, Taal E, Seydel E, van de Laar M. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *Journal of Medical Internet Research*. 2008; 10(2): e18.
103. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. *American Journal of Psychiatry*. 2002; 159(12): 2062-8.
104. Powell J, McCarthy N, Eysenbach G. Cross-sectional survey of users of Internet depression communities. *BMC Psychiatry*. 2003; 3(1): 19.
105. Lawlor A, Kirakowski J. Online support groups for mental health: A space for challenging self-stigma or a means of social avoidance? *Computers in Human Behavior*. 2014; 32: 152-61.
106. Muncer S, Burrows R, Pleace N, Loader B, Nettleton S. Births, deaths, sex and marriage... but very few presents? A case study of social support in cyberspace. *Critical Public Health*. 2000; 10(1): 1-18.
107. Macias W, Lewis LS, Smith TL. Health-related message boards/chat rooms on the Web: discussion content and implications for pharmaceutical sponsorships. *Journal of Health Communication*. 2005; 10(3): 209-23.
108. Griffiths KM, Reynolds J, Vassallo S. An online, moderated peer-to-peer support bulletin board for depression: User-perceived advantages and disadvantages. *JMIR Mental Health*. 2015; 2(2): e14.
109. Nimrod G. The membership life cycle in online support groups. *International Journal of Communication*. 2012; 6: 23.
110. Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A. User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study. *Internet Interventions*. 2017; 7: 9-15.
111. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM. Community structure of a mental health Internet support group: Modularity in user thread participation. *JMIR Mental Health*. 2016; 3(2): e20.

Chapter 2

Foreword

Chapter 2 comprises a systematic review of studies that have investigated participation in an online health community.

Initially, the purpose of this study was to identify what is and what is not currently known about active participation in MHISGs. However, during the early and informal stages of conducting the systematic review, it became clear that there was little research pertaining to MHISGs specifically. This prompted the expansion of the scope of the review to ISGs for any health condition with the objective of investigating similarities and differences among participation styles across and within support groups for different health conditions and to provide a context for the subsequent empirical investigation of MHISGs in this thesis.

The review was designed to address the following question:

RQ 1: In what ways has participation previously been defined and measured in an MHISG, and other online health communities?

The review was limited to studies which demarcated a sub-group of users by one or more metrics (referred to as styles of participation). Studies that employed qualitative analyses were included provided that they employed a method of categorising users that could be operationalised by other researchers.

The direction of the empirical studies in this thesis was informed by the two salient findings of the review. These were that:

- (i) active participation has been classified in many different ways. For example, machine-learning classifiers have been developed to identify 'leaders' and 'influential users', and algorithms have been designed to identify an optimal subset of 'key players' for maximum dissemination of information across a network; and

- (ii) active participation has most commonly been classified by level of engagement as operationalised by posting frequency, often with the assumption that higher posting frequency defines users of the greatest value to the community.

Both findings have implications for future MHISG research. First, they demonstrate that there are many promising means for classifying and investigating participation in ISGs which have not yet been applied to the field of mental health. Secondly, although post frequency has been the most commonly used metric of engagement, little is known about the factors associated with this metric. For example, prior to the empirical studies in this thesis, it was not known if high- and low-frequency MHISG posters differ in their personal characteristics, patterns of social network connectivity and biases in the type of content they contribute, or if the contributions of the two groups are characterised by the same style of engagement and differ only in quantity. Given that such evidence may improve our understanding of the utility of post frequency as a metric, the empirical studies in Chapters 4-7 were designed to investigate this issue.

Original Paper

From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities

Bradley Carron-Arthur¹, BPsych (Hons); Kathina Ali¹, MSc (Psych); John Alastair Cunningham^{1,2}, PhD; Kathleen Margaret Griffiths¹, PhD

¹National Institute for Mental Health Research, Research School of Population Health, The Australian National University, Acton, Australia

²Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:

Bradley Carron-Arthur, BPsych (Hons)
National Institute for Mental Health Research
Research School of Population Health
The Australian National University
63 Eggleston Road
Acton, 2601
Australia
Phone: 61 61256825
Fax: 61 61250733
Email: Bradley.Carron-Arthur@anu.edu.au

Abstract

Background: Understanding how people participate in and contribute to online health communities (OHCs) is useful knowledge in multiple domains. It is helpful for community managers in developing strategies for building community, for organizations in disseminating information about health interventions, and for researchers in understanding the social dynamics of peer support.

Objective: We sought to determine if any patterns were apparent in the nature of user participation across online health communities.

Methods: The current study involved a systematic review of all studies that have investigated the nature of participation in an online health community and have provided a quantifiable method for categorizing a person based on their participation style. A systematic search yielded 20 papers.

Results: Participatory styles were classified as either multidimensional (based on multiple metrics) or unidimensional (based on one metric). With respect to the multidimensional category, a total of 41 different participation styles were identified ranging from Influential Users who were leaders on the board to Topic-Focused Responders who focused on a specific topic and tended to respond to rather than initiate posts. However, there was little overlap in participation styles identified both across OHCs for different health conditions and within OHCs for specific health conditions. Five of the 41 styles emerged in more than one study (Hubs, Authorities, Facilitators, Prime Givers, and Discussants), but the remainder were reported in only one study. The focus of the unidimensional studies was on level of engagement and particularly on high-engaged users. Eight different metrics were used to evaluate level of engagement with the greatest focus on frequency of posts.

Conclusions: With the exception of high-engaged users based on high post frequency, the current review found little evidence for consistent participatory styles across different health communities. However, this area of research is in its infancy, with most of the studies included in the review being published in the last 2 years. Nevertheless, the review delivers a nomenclature for OHC participation styles and metrics and discusses important methodological issues that will provide a basis for future comparative research in the area. Further studies are required to systematically investigate a range of participatory styles, to investigate their association with different types of online health communities and to determine the contribution of different participatory styles within and across online health communities.

(*J Med Internet Res* 2015;17(12):e271) doi:[10.2196/jmir.4705](https://doi.org/10.2196/jmir.4705)

KEYWORDS

online health community; participation style; social network; participation inequality; systematic review

Introduction

Participation rates of people in online communities are known to be highly variable with some people contributing much more than others. Across all types of online communities, the variability in degree of user participation consistently follows a pattern [1]. In particular, this pattern in participation is described by a power law. This power law means, for example, that the top 1% of participants contribute as much as 75% of the posts in an online health community (OHC) [2,3]. This pattern is indicative of a coherent community [2], and these highly engaged individuals are repeatedly observed in well-established OHCs [4]. These individuals are of interest. Their high participation rates and predictable presence suggest that they may be of particular value to the OHC.

Although post frequency may constitute a simple indicator of engagement, from post frequency alone it is not possible to ascertain exactly what ways a person contributes. Post frequency does not indicate whether a person starts new discussions, welcomes newcomers, is available at critical times in the day when people are most likely to need support, or is knowledgeable about certain topics. In order to ascertain whether people contribute these different kinds of value, it is necessary to measure their participation based on various other metrics.

There may be value for those who are involved in the development of an OHC to identify users who contribute particular types of value to the OHC. This points to the need for multiple metrics to define user contributions. For example, in a qualitative paper on building and sustaining OHCs, Young described how certain core members were vital to the development and sustainability of an OHC [5]. As the community manager from the inception of this OHC, Young was able to provide an account of the different ways that these users had contributed to the development of the OHC including facilitating discussion and fostering a supportive culture. Young also suggested ways that OHC managers might harness the contributions of these individuals to help build the community by, for example, highlighting their best posts or inviting them to contribute to a community resource such as a newsletter.

For a variety of reasons, including time constraints and size of the community, not all community managers are able to have a strong qualitative understanding of the roles of particular individuals in their OHC. However, community managers would potentially benefit from a simple operationalization of user participation in terms of metrics that are automatically collected in the log data of the OHC software. This would help them to identify the core members and various other users who contribute in different ways so that they may apply the community building techniques recommended by Young [5].

OHCs also provide an opportune setting for interventions that encourage certain positive health behaviors [6]. Knowing who the most influential people are in an OHC, or how to reach most of the community via the smallest subset of people, might inform dissemination activities such as promoting new evidence-based treatments or recommending correct use of certain medications.

Finally, there is scientific value in investigating the ways in which different people participate in OHCs across multiple contexts. There may be patterns in the way in which people participate that can be found across multiple different OHCs. These patterns may help us learn more about the social dynamics of OHCs and the way that people seek help and provide it to others.

User profiling by categorizing participation styles is conducted in studies of online communities more broadly. There are some roles such as “newbies” and “celebrities” that may be found in any online community, but most others are likely to be specific to the type of community [7]. For example, “technical editors” and “substantive experts” are found in Wikipedia [8], but these may not be relevant to or found in OHCs. We expect that OHCs will have high-profile users who are akin to “celebrities,” but the nomenclature and the metrics used to define these users may be tailored to the supportive context and health discussion focus of the community. There may be further similarities and differences between participation styles in communities of different health types.

This study seeks to advance this area by conducting a systematic review of all studies that provide replicable, quantifiable criteria for categorizing the nature of participation in an OHC. We aimed to document all participation styles that had been identified to date and the OHCs from which they came. Our objective was to determine if any patterns were apparent in the nature of user participation across OHCs for different health conditions or within each.

Methods

A systematic review was conducted to identify articles that investigated participation styles in an online health community. For the current purposes, an online health community was defined as any Internet-based platform designed to enable people to communicate about health issues. A participation style was defined as any type of engagement with an OHC that can be measured quantitatively. This does not include simply the presence or absence of participation (ie, posters and lurkers), as this has been well documented elsewhere [9], but rather is aimed at understanding the nature of participation for those who are actively engaged in the community.

Search Strategy

Three databases (PubMed, PsycINFO, and Cochrane) were searched for all articles prior to December 2014. Adapted search terms from Eysenbach et al [10] and Griffiths et al [11] were used to identify the concept of OHC (see [Multimedia Appendix 1](#)). These search terms were combined with the following terms to identify the participation style concept: (participatory pattern*) OR (posting pattern*) OR (posting behavior pattern*) OR (use pattern*) OR (communication pattern*) OR (usage pattern*) OR (system use*) OR (traffic) OR (participative stance*) OR (participant contribution*) OR (posting habits*) OR (participation rate*) OR (posting rate*) OR (user engagement) OR (level* of engagement*) OR (pattern* of engagement*) OR (type* of engagement) OR (share information) OR (community structure) OR (social dynamics).

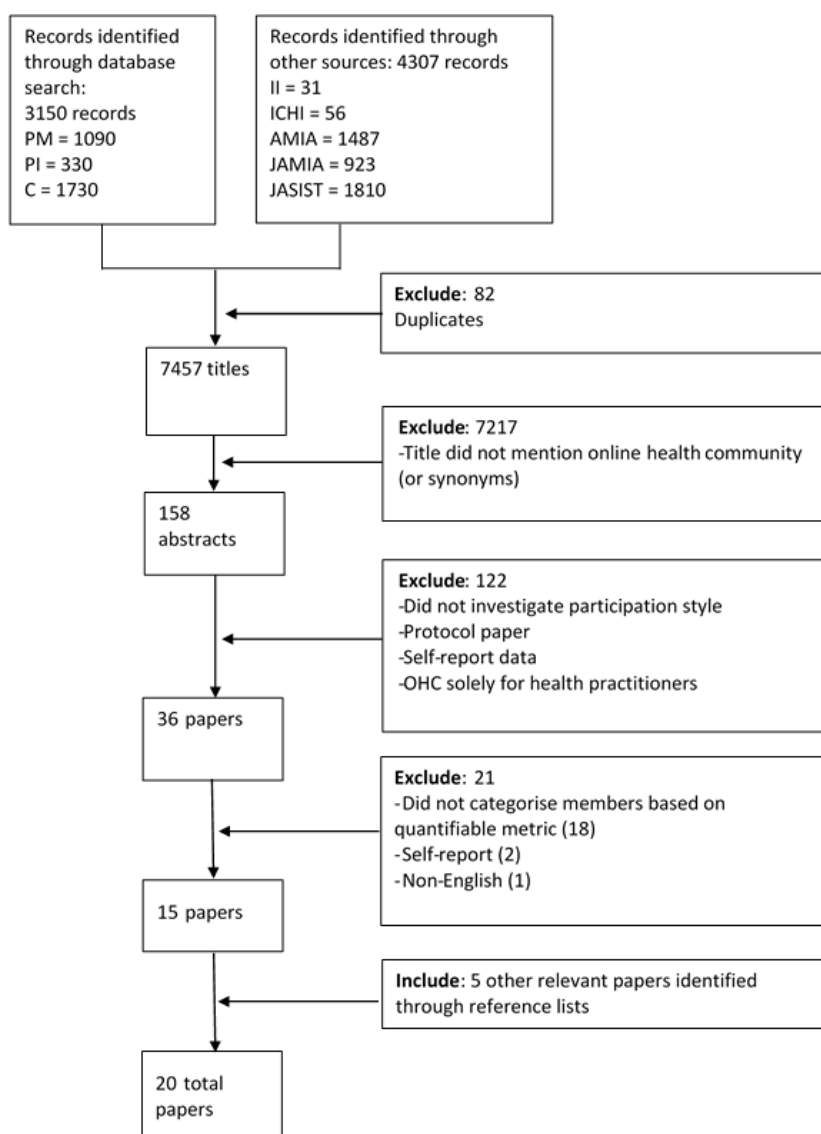
In addition, papers from relevant journals and conference proceedings in the computer and information science field published since 2005 (including the *American Medical Informatics Association Annual Symposium*, *Journal of the American Medical Informatics Association*, *Journal of the Association for Information Science and Technology*, and *International Conference on Healthcare Informatics*) and a new journal that was not yet indexed at the time of the search (*Internet Interventions*) were screened for relevant articles.

Article Selection

A total of 7457 articles were screened. Of these, 3150 were retrieved from the database search and 4307 were from the additional journals and conference proceedings. A total of 82

duplicate articles were identified and removed. Relevant articles were selected through a multistage process (Figure 1). Initially, titles were screened by 2 raters (BC and KA). Any article that mentioned an online community or synonym thereof in the title (or online health community in the case of the *Journal of the Association for Information Science and Technology*) was included. This reduced the number of articles to 158. The abstracts of these articles were subsequently screened by the 2 raters. Any article that investigated ways that people participate in an online health community was included. Articles based on self-report measures of OHC use and research protocols were excluded. The full articles for the 36 remaining abstracts were retrieved and read by both raters. Any disagreements between the raters were resolved by discussion.

Figure 1. Study identification flow diagram: PubMed (PM), PsychINFO (PI), Cochrane (C), Internet Interventions (II), International Conference on Healthcare Informatics (ICHI), American Medical Informatics Association Annual Symposium (AMIA), Journal of the American Medical Informatics Association (JAMIA), Journal of the Association for Information Science and Technology (JASIST).



Inclusion Criteria

The final set of articles included any study that (1) quantitatively investigated ways that people participate in an online health community, and (2) categorized users based on any quantifiable

metric that can be used to show they have engaged with the community.

Exclusion Criteria

Studies that converted written content to quantitative data by a means that was computerized (eg, machine learning algorithm) were included, but studies that relied on human interpretation of written content to create quantitative data were not. This ensured that the methods identified could be accurately replicated and would be scalable to large OHCs. For similar reasons, studies that used self-report data from surveys were not included. This meant that only studies reporting data that had been automatically logged by the OHC software or that had been extracted by programs that crawl publicly available data were included in this systematic review. Protocol papers, articles not written in English, and papers on OHCs solely for health practitioners were not included.

After applying these criteria, a set of 15 papers were included. The reference lists of included papers and those that cited them (as per Google Scholar) were hand searched. This yielded an additional 5 papers, resulting in a final set of 20 included papers.

Coding

The included papers were coded by 1 rater (BC). Each participation style identified by a paper was listed. Three attributes of each participation style were coded: (1) the name used by the authors to describe the participation style, for example, "superuser," (2) the metrics used to quantitatively describe their style of participation, for example, frequency of posts, and (3) the inclusion criteria used to determine who was categorized as having that participation style, for example, the top 1% of users whose frequency of posts was greatest were deemed to be superusers.

Results

Across the final set of 20 papers, users were categorized into participation styles a total of 74 times, of which 28 were duplicates. These duplicates included participation styles that had been assigned different names by different studies but used the same metrics and same inclusion criteria (or very similar) to define them. By merging all these redundant categorizations into the same participation style, we determined that 44 participation styles had been identified in OHCs to date.

[Table 1 \[2,3,12-29\]](#) shows a summary of information about the OHCs where the participation styles were identified. Some studies investigated more than one OHC. In total, there were 26 different OHCs. These were used for a variety of different health topics including smoking cessation (n=7), cancer (n=6), mental health issues (n=6), diabetes (n=5), multiple sclerosis (n=1), and social innovation in health care (n=1). These OHCs were based in different countries including the United States (n=8), Canada (n=2), Australia (n=1), Germany (n=1), New Zealand (n=1), Norway (n=1), Taiwan (n=1), and the United Kingdom (n=1). The country of origin for 10 OHCs was not reported. The sample of people drawn from each OHC ranged in size from 77 to 49,552 people. Most included between 1000 and 10,000 people; however, one group of 5 OHCs included more than 140,000 people between them. All of the studies were published in 2007 or later, with 12 of the 20 published since 2013.

[Table 2 \[2,3,12-29\]](#) shows a summary of these types of participation. Within [Table 2](#), we have grouped participation styles first into two categories: those based on multiple metrics (multidimensional) and those based on one metric (unidimensional). Each of these is also then divided into up to 3 categories according to the predominant type of metric used to define the participation style: activity-based, network-based, and content-based metrics. [Table 3 \[30,31\]](#) contains a list of the metrics and a description of what they measure.

There were 41 participation styles in the multidimensional category (13 activity based, 11 network based, and 17 content based). In all instances where a unidimensional participation style was identified, the studies divided the users into no more than 3 groups that we have summarized as high, medium, and low engagement. There were 8 different metrics used in the high engagement category (5 activity based, 3 network based), 3 in the medium category (2 activity based, 1 network based), and 4 in the low category (3 activity based, 1 network based).

The results of each subcategory of participation style (content based, network based, and activity based) are described in turn for the 41 multidimensional participation styles. Following this, the results of the unidimensional participation styles are described together for each of the 3 participation styles identified.

Table 1. Summary of online health community characteristics.

Online health community name	Year of study	Health condition	Country	Sample size, n
SOL-Cancer Forum	2007 [27]	Cancer	Not reported	84
Cancer Survivors Network	2014 [22,23]	Cancer	United States	27,173
Cancer Compass	2011 [28]	Cancer	United States	7991
WebChoice	2013 [13]	Cancer (breast and prostate)	Norway	103
Breastcancer.org	2014 [29]	Cancer (breast)	United States	49,552
Cancer Compass	2010 [17]	Cancer (melanoma)	United States	851
Five unnamed forums in English and Spanish	2013 [14]	Diabetes	Not reported	>140,000
BlueBoard	2014 [2]	Mental health	Australia	2932
DepressionCenter	2014 [3]	Mental health (depression)	Not reported	5151
PanicCenter	2014 [3]	Mental health (panic disorder)	Not reported	11,372
AlcoholHelpCenter	2014 [3]	Mental health (problem drinking)	Not reported	2597
PTT.CC—Psychosis Support Group	2009 [26]	Mental health (psychosis)	Taiwan	438
SharpTalk	2011 [19]	Mental health (self-harm)	United Kingdom	77
Deutsche Multiple Sklerose Gesellschaft	2014 [20]	Multiple sclerosis	Germany	1169
The Canadian Cancer Society's Smokers' Helpline Online	2012 [21]	Smoking	Canada	1670
QuitBlogs	2014 [18]	Smoking	New Zealand	3448
Alt.Support.Stop-Smoking	2014 [25]	Smoking	Not reported	8236
QuitPlan	2008 [12]	Smoking	United States	233
QuitNet	2010 [15]	Smoking	United States	7569
StopSmokingCenter	2013 [16]	Smoking	United States	Not reported
	2012 [21]			1627
	2014 [3]			44,870
#HCSMCA	2013 [24]	Social innovation in health care	Canada	486

Table 2. Summary of participation styles including name, metrics, and inclusion criteria.

Name	Metrics	Inclusion criteria
Multidimensional		
Content based		
Influential user [23]	69 activity, network, and content features including influential responding replies	Machine learning classifier (relying initially on expert judgement to identify exemplars)
Leader [22]	68 activity, network and content features	Machine learning classifier (relying initially on expert judgement to identify exemplars)
Opinion leader [16]	Word vectors, degree	Latent semantic analysis and high degree
Information providers [29]	Social support type	High information support
Community builders [29]	Social support type	High companionship support
Emotional support providers [29]	Social support type	High emotional support
Information seekers [29]	Social support type	High information support seeking
Emotional support seekers [29]	Social support type	High emotional support seeking
Information enthusiasts [29]	Social support type	High information support seeking, high information support
All-around contributors [29]	Social support type	No particular metric stands out
Balanced source user [20]	Source of information	Cited information from a range of sources
Social media fan [20]	Source of information	High social media
Organization follower [20]	Source of information	High organizations
Homepage promoter [20]	Source of information	High static informational websites
Seeker of health care [20]	Source of information	High health practitioners
User of uncommon sources [20]	Source of information	High uncommon sources
Sophisticated contributor [20]	Word count, source of information	High word count, high academic references
Network based		
Key player [15]	Degree (nonredundant)	Key Player 1.4 software
Hub [14,17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Authority [14,17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Facilitator [17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Trusted user [14]	Out-degree, in-degree	PageRank algorithm
Help-seeker [14]	Out-degree, in-degree	Low in-degree, high out-degree (within the scope of the edge between 2 users)
Star [27]	Out-degree, in-degree	Top ranked individual (outlier)
Prime givers [14,27]	Out-degree, in-degree	Very high out-degree, high in-degree
Serious members [27]	Out-degree, in-degree	Moderate out-degree, moderate in-degree
Moderate users [27]	Out-degree, in-degree	Low out-degree, low in-degree
Takers [27]	Out-degree, in-degree	No out-degree, low in-degree
Activity based		
Caretaker [19]	Time logged in, episodes, reading, posting, thread initiation	High time logged in, low episodes, high reading, low posting, low thread initiation
Here for you [19]	Thread initiation, posting, forum	Low thread initiation, high posting in support forum
Butterfly [19]	Time logged in, episodes, posting, forum	High time logged in, high episodes, high posting in support forum

Name	Metrics	Inclusion criteria
Crisis-oriented individual [19]	Posting, forum	High posting in support forum
Discussant [19,20]	Thread initiation, posting, forum	High thread initiation, high posting in discussion forum
Average user [20]	Thread initiation, posting, forum, topic, days active, word count, source of information	No particular metric stands out
Highly active relational poster [20]	Posts per day, thread participation, thread initiation	High posts per day, high thread participation, low thread initiation
Topic-focused responder [20]	Thread initiation, posting, topic, days active	Low thread initiation, low posts per day, high fraction of topic-related posts, low days active
Topic-spammer [20]	Posting, days active, word count, topic, source of information	Low days active, high posting, low word count, high fraction of topic-related posts, low references
Long-term high-activity users [25]	Days active, posting	High days active, high posting
Short-term high-activity users [25]	Days active, posting	Low days active, high posting
Short-term low-activity users [25]	Days active, posting	Low days active, low posting
Long-term low-activity users [25]	Days active, posting	High days active, low posting
Unidimensional		
Activity based		
High-engaged user	Posting	>2 posts [12]; top 1% of users [2,3]; top 10 users [24]; >180 posts [18]; top 100 users [21]
	Reading	>5 posts [12]
	Time logged in	Top 33.3% of users [13]
	Thread initiation	Top 100 users [21]
	Thread participation	Top 100 users [21]
	Network based	
	Friendship	Mutual friend nomination between 2 users and >4 interactions between them [15]
Moderate-engaged user	In-degree	Top 10 users [24]; high in-degree [26]
	Out-degree	Top 10 users [24]; high out-degree [26]
	Posting	2-10 percentile (9%) of users [2,3]
Low-engaged user	Time logged in	Middle 33.3% of users [13]
	Network based	
	Friendship	Friend nomination of another user and >0 interactions with them [15]
Low-engaged user	Posting	1-2 post [12,18]; bottom 90% of users [2,3]
	Reading	1-5 posts [12]
	Time logged in	Bottom 33.3% of registered users [13]
	Network based	
	Friendship	Any interactions with another user [15]

Table 3. A description of the metrics used to classify participation styles.

Metric	Description
Activity-based metrics: measure the individual actions taken by users in an OHC	
Posting	Number of posts a person has made in the OHC
Time logged in	Amount of time a person has spent accessing the OHC
Reading	Number of posts that a person has read
Thread initiation	Number of times a person has created a thread
Episodes	Number of times a person has accessed the OHC
Days active	Number of days between a person's first and last post
Forum	Number of posts a person has made in a particular subforum of the OHC, eg, support or discussion
Thread participation	Number of different threads a person has posted in
Network-based metrics: measure the relationship and interactions between users	
Degree (in/out)	The number of people a person has communicated with. Where it is possible to tell who the source of the communication was and to whom it was directed, the number of people a person has made outgoing communication with is called the "out-degree" and the number of people that a person has received communication from is called the "in-degree." When it is not possible to tell the direction, the communication is counted for both people as a measure of degree. Degree is considered to be a measure of a user's centrality in a network [30,31].
Friendship	The extent to which a person is connected with at least one other person in the OHC as defined by 3 thresholds: Low—any interactions with another user; Moderate—friend nomination of another user and >0 interactions with them; and High—mutual friend nomination between 2 users and >4 interactions between them.
Content-based metrics: measure the nature of the content within posts	
Word vectors	A representation of the proportion of words in a message that fit a certain topic.
Influential Responding Replies	Number of posts a person has made that have influenced the sentiment of the thread initiator
Social support type	Number of posts a person has made that either provide or seek information support, emotional support, or companionship
Topic	Number of posts a person has made which included subject matter on a specific topic
Source of information	Number of citations a person has made from a particular source
Word count	Number of words in a post

Multidimensional

Content-Based

Leaders and Influential Users

Zhao et al [23] created a machine learning classifier with 69 metrics that was used to identify influential users in an OHC. These users were regarded as leaders who could influence the emotional sentiment of other users. This study built on previous research by Zhao et al [22], which used 68 metrics such as number of posts, in-degree, and days active in a classifier to first identify leaders in the OHC. Zhao et al [23] then created a metric called "influential responding replies (IRRs)." This was the number of times a person was able to affect the sentiment of another person when responding to their initial post. It was found that this metric alone outperformed the classifier with 68 metrics, and together they created the best performing classifier. In order to train this IRR-enhanced classifier, it was necessary to have a list of users who were deemed to be influential users by moderators of the OHC. There were 41 users in this list. In total, the moderators identified 126 influential users. A list of the top 50, 100, and 150 influential users identified by the classifier was made up with 90%, 77.7%, and 68.7% users from the moderator list of 126 respectively. The highest percentage

possible in the 150 influential user condition was 84.0% (126/150).

Opinion Leaders

Myeni et al [16] used latent semantic analysis to identify users who were involved in discussion about particular concepts such as personal experiences, advice, or adherence to interventions. Users whose mean word vector scores for a concept were one standard deviation above the sample mean were grouped together in a social network. Within each theme-based social network, Opinion Leaders were identified as people who had the highest degree. These people were considered to be influential in their specific domain and may be particularly useful to identify when delivering relevant targeted interventions. Subsequent research has shown that exposure to users who were abstinent from smoking in the theme-based networks of "social support" and "traditions" were more likely to be abstinent themselves [32].

Information Providers, Emotional Support Providers, Information Support Seekers, Emotional Support Seekers,

Community Builders, Information Enthusiasts, and All-Around Contributors

Wang et al [29] created a machine learning classifier to determine which posts in a cancer OHC with more than 2.8 million posts contained each of the following types of content: providing informational support, providing emotional support, seeking informational support, and seeking emotional support and companionship. The authors then used a k-means clustering algorithm [33] to categorize all users based on the proportion of posts they made with each type of content. This produced 7 types of users. Five were typified by writing a high proportion of posts that predominantly contained one type of each of the above 5 content types. The remaining two, information enthusiasts and all-around contributors, were typified by having equally high proportions of posts seeking and providing informational support, and having equal amounts of all types of content, respectively. The all-around contributor was the most common type of user of all 7 (making up 32% of all users). Community builders were among the least common (8%) but were responsible for writing the most posts on average along with all-around contributors. Those who primarily engaged in informational and emotional support types posted less and did not remain in the forum as long as community builders and all-around contributors.

Balanced Source User, Social Media Fan, Organization Follower, Homepage Promoter, Seeker of Health Care, and User of Uncommon Sources

Sudau et al [20] observed that people tend to favor different sources of information to support the points that they make in posts. A number of participation styles represent this bias. In order to determine these participation styles, Sudau et al used a k-means clustering algorithm [33] to form 6 groups of similar users based on the frequency of different hyperlinks they used from 8 domain classes. The groups were labeled according to what Sudau et al thought best described their referencing tendencies.

Sophisticated Contributor

A sophisticated contributor is a user whose posts are longer than those of the average user participation style and contain more references. In contrast to the activities of most users, these references are more often to scientific publications than to social media sources. Sudau et al [20] identified this participation style in 4 of 171 users. Sophisticated contributor posts were three times as long and contained five times as many references as posts by Average Users.

Network Based

Key Players

Cobb et al [15] sought to identify a set of users who were maximally connected to other users throughout the social network of the OHC. A set of key players is a small group of a specified number of users who are connected with as many other people in the network as possible, for example, through private message, posting, or friendship. Cobb et al used Key Player 1.4 software [34] to determine the reach of a set of 50 key players. These 50 key players were connected to 64% of other users in the network. Note that these are not necessarily the 50 most

connected individuals in the OHC; that is, they are not the top 50 users ranked by degree. Rather, the algorithm considers redundancy. If introducing a new key player to the set does not increase the set's overall reach, that player is not added. The optimum key player set of 50 users may not necessarily contain all the users in the 49 set nor will either necessarily contain the user who, as an individual, is the most connected person in the network. The intention of the algorithm is to enable maximum access to the whole network through minimal nodes. This, for example, enables maximum efficiency in dissemination of information.

Hubs, Authorities, and Facilitators

Hubs and authorities are concepts borrowed from the computer science literature on the Web. Hubs and authorities are identified using the hyperlink-induced topic search (HITS) algorithm [35]. In this algorithm, every website receives both a hub and an authority score. High-scoring authorities are websites that are linked to high-scoring hubs. High-scoring hubs are websites that link to high-scoring authorities. Websites with high authority scores tend to those that provide good information on a specific topic. Hubs direct people to these various authorities. The algorithm can be applied to any network consisting of nodes and links between them by analyzing the pattern of out-degree and in-degree across the network. Accordingly, both Chomutare et al [14] and Durant et al [17,28] have used the HITS algorithm to identify people in OHCs as authorities and hubs. The 3 papers have posited that those identified as hubs are people who disseminate information by promoting discussion. They have a relatively high out-degree in the network compared with their in-degree. They are important for sustaining the activity levels of the community. Authorities are people whose opinion is highly respected in the community. They have a relatively high in-degree. A third participation style—a facilitator—was also proposed by Durant et al [17]. A facilitator is a person who is ranked similarly highly as a hub and as an authority. They are considered to be more effective for sustaining communication in the OHC than those who are hubs or authorities alone. Durant et al [28] sought to track the presence of facilitators over time by segmenting and analyzing the network each year over an 8-year period and found that the top 5% were rarely the same individuals in consecutive years.

Trusted Users

Similar to the HITS algorithm, the PageRank algorithm [36] is another method originating in the computer science literature on the Web. Rather than identifying 2 types of users, the PageRank algorithm identifies one type. The score given to each node in the network by the PageRank algorithm is the probability of arriving at that node given a random walk around the network via the links between them. This means that nodes that are linked to more often have higher probabilities of being landed on, and nodes that are linked to more often by other high-scoring nodes have even higher scores. If it is assumed that a directional link between 2 nodes is a vote of support to the other, this algorithm identifies trusted users. This algorithm was the basis for Google search. Chomutare et al [14] have applied it to an OHC and have made the same assumption. They found that 6 out of 10 of the highest ranked users by in-degree were also in the top 10 identified by the PageRank algorithm.

Help-Seekers

In a relationship between 2 people where one communicates with the other much more often, the person who instigates more communication (higher out-degree than in-degree) is labeled a Help-Seeker. Chomutare et al [14] suggested that this pattern of metrics might reflect a person who is struggling with their health issue. However, the authors noted that the nature of the help-seeking is not exactly clear as the user may either be strongly motivated to engage in self-care or they may be a particularly needy user, and neither can be concluded without content analysis. The authors originally suggested the label “needy user” for this participation style, but we have renamed it “help-seeker” given the ambiguity and lack of clarity around the concept of needy in this context.

Star, Prime Givers, Serious Members, Moderate Users, and Takers

The earliest recorded participation styles were identified by Bambina [27] who compared the in-degree and out-degree of users and grouped them around a pattern in the results that was related to engagement. Bambina first noted one outlier: a person who had both the highest in-degree and out-degree. Bambina referred to this person as the “star.” This person provided the most support to others including notably many new individuals with whom many others did not communicate. Bambina noted that the next most engaged people by both in-degree and out-degree all tended to provide more support than they received, that is, have higher out-degree than in-degree. These were named “prime givers” (n=6). Chomutare et al [14] observed the same pattern in a social network analysis that they conducted, but they did not report whether it was associated with providing support. Bambina also noted 2 groups who had relatively similar in-degree and out-degree within each group. These were the designated “serious members” (10) and “moderate users” (n=15). Last was a group labeled the “takers” who never provided support but who initiated a conversation and received support from others (n=52).

Activity-Based

Caretakers

Jones et al [19] identified one user in a sample of 77 people as having a participation style called the “caretaker.” They identified this person, as they did for all participation styles, through visual inspection of scatterplots of various metrics. The OHC was a support group for young people who self-harm. Given the large amount of time the person spent logged in, they actively participated very little. The times they did post were largely in response to other users rather than initiating their own threads. Jones et al concluded that this person might be watching over the whole forum and looking out for others in need. This person undertook the caretaker role despite the OHC being a moderated forum.

Here for You

One user in a sample of 77 people was considered to take the “here for you” participation style by Jones et al [19]. Like the caretaker, they did not create many threads of their own. However, in contrast to the latter, they did post large amounts of comments in response to other people who needed support.

Butterfly

Another user in the Jones et al [19] sample was classified as being characterized by a butterfly participation style. This person logged in many more times than anyone else. They spent short amounts of time checking out a few pages and then logged out again. They posted mostly in the support forum (as opposed to the discussion forum or off-topic forum). Like the crisis-oriented individuals in the following section, they were considered by the moderators to be in crisis and needing support as opposed to providing it.

Crisis-Oriented Individuals

Six users of the Jones et al [19] sample were classified as crisis-oriented in their participation style. These people posted in large numbers in the support forum. It is not possible to confirm from the objective metrics alone whether such people were in crisis or providing support; however, it was confirmed by the moderators of the forum that all 6 were indeed in crisis. These users did not visit the OHC as frequently as the user with the butterfly participation style.

Discussants

A discussant is a user who is mainly focused on discussion about health-related topics as opposed to providing or receiving support. They initiate a high number of threads in the discussion section of the OHC and participate actively in them. This participation style was identified by both Jones et al [19] and Sudau et al [20].

Average Users

A user type that is not distinctly based on any metric, the average user category was identified by the application of a second k-means clustering algorithm conducted by Sudau et al [20]. This analysis was designed to form 6 groups of similar users based on 9 metrics that measured their active participation in the community. Sudau et al labeled the groups according to their distinguishing features. Average users were a group that were thought not to exhibit any distinguishing features. This group constituted 63% of the people included in the analysis.

Highly Active Relational Posters

These are the most active users of an OHC by post frequency. Sudau et al [20] noted these users maintain “small talk,” which may be good for community building. They participate in many different threads but do not initiate many themselves.

Topic-Focused Responders

A user whose activity is concentrated on a specific topic, the topic-focused responder is distinct from a discussant in that they do not post as much and do not initiate as many threads. Sudau et al [20] included only people who had made at least five posts on a certain topic in their analysis. Topic-focused responders met this criterion but they did not have many other posts. They tended to focus mainly on responding to others who had initiated the topic. Sudau et al suggested this style may be similar to the here for you participation style identified by Jones et al [19], but we have separated them because of the distinction between discussion and support.

Topic-Spammers

This is a user who is active for a very short period, that is, only a few days. In that time, they contribute a high number of posts on a specific topic in the discussion forum. However, these are not particularly sophisticated posts, rather they are short and lack references. This participation style was identified by Sudau et al [20].

Short-Term and Long-Term, High-Activity and Low-Activity Users

Stearns et al [25] noted that the bulk of users in a smoking cessation OHC are made up of short-term users (active for approximately less than 1 week), who, regardless of whether they have high or low activity, tend to be involved in the OHC for personal gain. Long-term users with low activity are noted to have smaller social circles and a stronger interest in particular topics. Stearns et al state that long-term high-activity users are most like Young's [5] "core members" who are vital to the sustainability of the OHC.

Unidimensional

All but one of the studies [24] that made unidimensional classifications did so for the purpose of determining if the type and level of engagement a person showed was predicted by demographic factors and whether high engagement predicted specific health outcomes. Some studies made statements about the nature of participation of users in the OHC. Given that the purpose of this review was to investigate the nature of participation, we focus on reporting these findings in the following sections considering first high-engaged users, followed by moderate- and low-engaged users.

High-Engaged Users

All 8 studies that made a unidimensional categorization [2,3,12,13,15,18,21,24] classified users into a participation style that we call high-engaged users. There were 8 different metrics used across these studies that all indicate a different type of high engagement. These included posting frequency, thread initiation, thread participation, level of in-degree/out-degree, reading of posts, time logged in, and friendship (see Table 3 for definitions).

Frequency of posting was the most commonly used metric used by 6 of the 8 studies [2,3,12,18,21,24]. It was used to classify users in a total of 9 OHCs, with 4 of those being for smoking cessation, 4 for mental health issues, and 1 for social innovation in health care. Users who were highly engaged according to posting frequency were regarded by all but one of the studies [12] as being valuable to the OHC because they sustained activity levels and in doing so facilitated the engagement of others. Four of the 6 studies referred to these people as either "superusers" [2,3,21] or "community leaders" [24]. This regard spanned across all the types of OHCs mentioned earlier.

Thread initiation and thread participation (together with posting frequency) were used by one study [21] to classify the top 100 ranked users, denoting them "superusers." The moderators of the OHC were asked to identify leaders within it. The authors noted that although most studies have previously identified leaders in an OHC using posting frequency alone, the moderators

thought it was necessary to also include users who start many conversations and who participate in many different conversations in their definition of a "superuser."

In-degree and out-degree were employed by 2 studies to classify users as highly engaged [24,26]. The authors of one study [24] regarded users with high in-degree (top 10) as authorities on topics, similar to the hubs and authorities discussed earlier. This study was conducted on an OHC that existed within Twitter. It was noted that those people with the highest in-degree were also people who had the highest number of followers on Twitter in general. Users with high in-degree were considered to be valuable for engaging other less active users in discussion. It was noted that the 6 users were both top 10 ranked users by in-degree and out-degree. These 6 people were thought to be communicating on topics that resonated with the community and were considered to be "community leaders." In a study of a mental health OHC for psychosis, Chang et al [26] referred to users with either a high in-degree or out-degree as "stars" after Bambina's [27] single outlying user.

Other metrics employed to classify users as highly engaged included reading [12], time logged in [13], and friendship [15].

Moderate-Engaged Users

Four studies classified users as moderately engaged based on 3 different metrics. Two were based on posting frequency [2,3], and one each on time logged in [13] and friendship [15].

Low-Engaged Users

Six studies classified users as low engaged based on 4 different metrics. Four were based on posting frequency [2,3,12,18], and one each on reading [12], time logged in [13], and friendship [15].

Discussion

Principal Findings

This systematic review synthesized findings from studies that investigated the nature of participation in an OHC by categorizing users based on metrics of participation. The aim of this review was to identify the different ways in which users participate and contribute to OHCs, although we acknowledge that the resultant list of participation styles may not provide a comprehensive account of all possible styles. Our objective was to determine whether any patterns were apparent in the types of participation styles that were identified across and within different health conditions. With the exception of an overlap in engagement measured by posting frequency (which has been discussed elsewhere [3]), there was little overlap in participation styles identified across OHCs for different health conditions or within OHCs for specific health conditions. Consequently, it is not possible for this study to address this objective. This area of research is in its infancy, with most of the studies included in this review being published in the last 2 years. Despite this shortcoming, the current review delivers a nomenclature for OHC participation styles and metrics that will provide a basis for future comparative research in the area. To inform future research, we discuss in the following section some

methodological considerations for studies seeking to replicate or expand on the methods identified by this review.

Methodological Considerations

Posting Frequency

It was common for studies to use posting frequency as the sole means of classifying highly engaged users in an OHC. It was also common among these studies for researchers to regard these users as being particularly valuable to the OHC. However, it is not possible to know from post frequency alone in what way a person is contributing to an OHC. They might be contributing trivial or critical messages or their post might in other ways fail to support others. The rationale for the inference that high engagement is synonymous with high value may relate to another commonality across papers. The authors in question were also community managers of the OHCs that they were studying; therefore, they may have based their conclusions on reading content posted by these users. However, content analysis research is required to investigate whether posting frequency is a valid means of identifying generically valuable users.

Machine Learning

Zhao et al [22,23] used a complex method of identifying the participation styles of leaders and influential users that may be subject to issues with generalizability. Ideally, the classifier would be transferable across OHCs. However, there is currently no evidence to support such transferability. Indeed by using 69 metrics in their machine learning classifier, they may have created a model that is overfitted to the data of the OHC from which it came and it may not work well at identifying leaders or influential users in other OHCs, even of the same health condition. Furthermore, an essential prerequisite for the development of the method was identifying a priori, using subjective judgments, a sample of leaders and influential users for use in the learning classifier trial. Thus, if Zhao's classifier is not generalizable, research to identify a new model requires expertise, or access to expertise, in identifying leaders and influential users through qualitative methods in addition to advanced understanding of machine learning methods. Despite these challenges, research in this area offers promise, particularly as influential users most closely resemble those vital users whom Young [5] described as core members. For those who are not inclined to build their own classifier, it is noteworthy that one particularly useful and generalizable aspect of the method for determining influential users was the discovery of the metric influential responding replies, which is defined as the number of posts a person has made that have influenced the sentiment of the thread initiator. Zhao reported that this metric was a better predictor of influential user status than the other 68 metrics combined. IRRs are determined by analyzing the degree of positive and/or negative sentiment expressed in the text. There are many existing programs that can conduct this kind of sentiment analysis, such as Linguistic Inquiry and Word Count [37]. However, note that it is important to test the validity of these programs in any new dataset by comparing human and computer ratings. As Zhao points out, the word "positive" in the context of a cancer diagnosis can be a negative concept. Applying a standard sentiment analysis program in this context would yield invalid results.

Wang et al [29] also used a machine learning classifier; however, their method may be more reliably replicated without expert knowledge. The classifier was designed to detect the presence or absence of certain types of social support in posts. They used 5 human coders to classify a sample of posts that could be used for training the classifier. These people were not domain experts. Similar research has involved contracting online Amazon Mechanical Turk workers to code the presence of social support in posts for the same purpose [38]. These people also did not have prior experience in this area.

Centrality Algorithms

Similar to IRR, some participation styles described users who were useful in a particular way that would be potentially identifiable in any OHC, or for that matter, any social network. These were based on algorithms that used measures of centrality such as in-degree and out-degree. This includes authorities, hubs, facilitators, and trusted users. While these categories are quite useful, it should be noted that these algorithms are calculated in such a way that they introduce bias based on time elapsed such that users who participate earlier in the OHC receive higher scores [39]. There are methods to adjust for this [40].

K-means Clustering and Multivariate Outliers

Other more specific participation styles described users who have particular characteristics and may be found only in a subset of OHCs. This included, for example, the caretaker or the topic-spammer. The techniques used to identify these participation styles, k-means clustering algorithms and multivariate outliers, may not necessarily identify the same participation styles in other OHCs. However, they may be useful for identifying other particular or unique ways of participating in OHCs.

Limitations and Future Research

The scope of this study is quite broad. We included all studies that categorized a type of participation in an OHC despite the possibility that the culture and nature of participation in populations with different health conditions and with or without moderators could differ markedly. There was little overlap in the use of categorizations to define particular participation styles either in OHCs broadly or within specific health conditions. Thus, it is not possible to draw many specific conclusions at this early stage. A possible limitation and reason for this is that we may not have included all relevant studies, as our search terms may not have encompassed all the different terms used to describe participation styles at this early stage of research. Nevertheless, by synthesizing the findings of the included studies, this review provides a basis for future research to investigate the validity of styles identified to date by attempting to replicate findings for specific OHCs and exploring their validity across different OHCs. Future research should also investigate new participation styles not documented in this review.

Conclusion

Our systematic review identified a range of participation styles. Some of them may be generalizable to other OHCs. Others were more specific to particular OHCs but were identified by methods

that could be used elsewhere. The findings of this review are intended to support the work of community managers in building community, organizations seeking to design targeted interventions and disseminate information through certain types of people in OHCs, and researchers seeking to understand the nature of peer support. We anticipate that this review will be

useful for these groups in conducting investigations to determine the presence of participation styles that may be relevant to their work. However, it is too early to draw any conclusions about which OHCs would be most likely to contain users who have specific participation styles.

Acknowledgments

BC is supported by an Australian Postgraduate Award. KA is supported by a Young and Well CRC PhD Scholarship. KG is supported by the Australian National Health and Medical Research Council (NHMRC) Research Fellowship .

Conflicts of Interest

None declared.

Multimedia Appendix 1

OHC concept search terms.

[[PDF File \(Adobe PDF File\), 17KB - jmir_v17i12e271_app1.pdf](#)]

References

1. Nielsen Norman Group. Participation inequality: lurkers vs contributors in internet communities URL: <http://www.nngroup.com/articles/participation-inequality/> [accessed 2014-06-06] [WebCite Cache ID 6Q7EwEncA]
2. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions* 2014 Oct;1(4):165-168. [doi: [10.1016/j.invent.2014.09.0033](https://doi.org/10.1016/j.invent.2014.09.0033)]
3. van Mierlo T. The 1% rule in four digital health social networks: an observational study. *J Med Internet Res* 2014;16(2):e33 [FREE Full text] [doi: [10.2196/jmir.2966](https://doi.org/10.2196/jmir.2966)] [Medline: [24496109](https://pubmed.ncbi.nlm.nih.gov/24496109/)]
4. van Mierlo T, Hyatt D, Ching A. Mapping Power Law Distributions in Digital Health Social Networks: Methods, Interpretations, and Practical Implications. *J Med Internet Res* 2015;17(6):e160 [FREE Full text] [doi: [10.2196/jmir.4297](https://doi.org/10.2196/jmir.4297)] [Medline: [26111790](https://pubmed.ncbi.nlm.nih.gov/26111790/)]
5. Young C. Community management that works: how to build and sustain a thriving online health community. *J Med Internet Res* 2013;15(6):e119 [FREE Full text] [doi: [10.2196/jmir.2501](https://doi.org/10.2196/jmir.2501)] [Medline: [23759312](https://pubmed.ncbi.nlm.nih.gov/23759312/)]
6. Brownstein CA, Brownstein JS, Williams DS, Wicks P, Heywood JA. The power of social networking in medicine. *Nat Biotechnol* 2009 Oct;27(10):888-890. [doi: [10.1038/nbt1009-888](https://doi.org/10.1038/nbt1009-888)] [Medline: [19816437](https://pubmed.ncbi.nlm.nih.gov/19816437/)]
7. Golder S, Donath J. Social roles in electronic communities. 2004 Presented at: Association of Internet Researchers (AoIR) Conference Internet Research 50; 2004; Brighton, England.
8. Wesler H, Cosley D, Kossinets G, Lin A, Dokshin F, Gay G. Finding social roles in Wikipedia. : ACM; 2011 Presented at: iConference; 2011; New York, NY.
9. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MJ. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *J Med Internet Res* 2008;10(2):e18 [FREE Full text] [doi: [10.2196/jmir.992](https://doi.org/10.2196/jmir.992)] [Medline: [18653442](https://pubmed.ncbi.nlm.nih.gov/18653442/)]
10. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ* 2004 May 15;328(7449):1166 [FREE Full text] [doi: [10.1136/bmj.328.7449.1166](https://doi.org/10.1136/bmj.328.7449.1166)] [Medline: [15142921](https://pubmed.ncbi.nlm.nih.gov/15142921/)]
11. Griffiths KM, Calcar AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? *J Med Internet Res* 2009;11(3):e40 [FREE Full text] [doi: [10.2196/jmir.1270](https://doi.org/10.2196/jmir.1270)] [Medline: [19793719](https://pubmed.ncbi.nlm.nih.gov/19793719/)]
12. An LC, Schillo BA, Saul JE, Wendling AH, Klatt CM, Berg CJ, et al. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. *J Med Internet Res* 2008;10(5):e55 [FREE Full text] [doi: [10.2196/jmir.1018](https://doi.org/10.2196/jmir.1018)] [Medline: [19103587](https://pubmed.ncbi.nlm.nih.gov/19103587/)]
13. Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM. How user characteristics affect use patterns in web-based illness management support for patients with breast and prostate cancer. *J Med Internet Res* 2013;15(3):e34 [FREE Full text] [doi: [10.2196/jmir.2285](https://doi.org/10.2196/jmir.2285)] [Medline: [23454601](https://pubmed.ncbi.nlm.nih.gov/23454601/)]
14. Chomutare T, Arsand E, Fernandez-Luque L, Lauritzen J, Hartvigsen G. Inferring community structure in healthcare forums. An empirical study. *Methods Inf Med* 2013;52(2):160-167. [doi: [10.3414/ME12-02-0003](https://doi.org/10.3414/ME12-02-0003)] [Medline: [23392282](https://pubmed.ncbi.nlm.nih.gov/23392282/)]
15. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. *Am J Public Health* 2010 Jul;100(7):1282-1289 [FREE Full text] [doi: [10.2105/AJPH.2009.165449](https://doi.org/10.2105/AJPH.2009.165449)] [Medline: [20466971](https://pubmed.ncbi.nlm.nih.gov/20466971/)]

16. Myneni S, Cobb NK, Cohen T. Finding meaning in social media: content-based social network analysis of QuitNet to identify new opportunities for health promotion. *Stud Health Technol Inform* 2013;192:807-811. [Medline: [23920669](#)]
17. Durant KT, McCray AT, Safran C. Social network analysis of an online melanoma discussion group. *AMIA Jt Summits Transl Sci Proc* 2010;2010:6-10 [FREE Full text] [Medline: [21347134](#)]
18. Healey B, Hoek J, Edwards R. Posting behaviour patterns in an online smoking cessation social network: implications for intervention design and development. *PLoS One* 2014;9(9):e106603 [FREE Full text] [doi: [10.1371/journal.pone.0106603](#)] [Medline: [25192174](#)]
19. Jones R, Sharkey S, Smithson J, Ford T, Emmens T, Hewis E, et al. Using metrics to describe the participative stances of members within discussion forums. *J Med Internet Res* 2011;13(1):e3 [FREE Full text] [doi: [10.2196/jmir.1591](#)] [Medline: [21239373](#)]
20. Sudau F, Friede T, Grabowski J, Koschack J, Makedonski P, Himmel W. Sources of information and behavioral patterns in online health forums: observational study. *J Med Internet Res* 2014;16(1):e10 [FREE Full text] [doi: [10.2196/jmir.2875](#)] [Medline: [24425598](#)]
21. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. *J Med Internet Res* 2012;14(3):e66 [FREE Full text] [doi: [10.2196/jmir.1854](#)] [Medline: [22732103](#)]
22. Zhao K, Greer G, Yen J, Mitra P, Portier K. Leader identification in an online health community for cancer survivors: a social network-based classification approach. *Inf Syst E-Bus Manage* 2014 Sep 16:1-17. [doi: [10.1007/s10257-014-0260-5](#)]
23. Zhao K, Yen J, Greer G, Qiu B, Mitra P, Portier K. Finding influential users of online health communities: a new metric based on sentiment influence. *J Am Med Inform Assoc* 2014 Oct;21(e2):e212-e218 [FREE Full text] [doi: [10.1136/amiajnl-2013-002282](#)] [Medline: [24449805](#)]
24. Gruzd A, Haythornthwaite C. Enabling community through social media. *J Med Internet Res* 2013;15(10):e248 [FREE Full text] [doi: [10.2196/jmir.2796](#)] [Medline: [24176835](#)]
25. Stearns M, Nambiar S, Nikolaev A, Semenov A, McIntosh S. Towards Evaluating and Enhancing the Reach of Online Health Forums for Smoking Cessation. *Netw Model Anal Health Inform Bioinform* 2014 Dec;3:-. [doi: [10.1007/s13721-014-0069-7](#)] [Medline: [26075158](#)]
26. Chang H. Online supportive interactions: Using a network approach to examine communication patterns within a psychosis social support group in Taiwan. *J Am Soc Inf Sci* 2009 Jul;60(7):1504-1517. [doi: [10.1002/asi.21070](#)]
27. Bambina A. *Online Social Support: The Interplay of Social Networks and Computer-Mediated Communication*. Youngstown, New York: Cambria Press; 2007.
28. Durant KT, McCray AT, Safran C. Identifying Temporal Changes and Topics that Promote Growth Within Online Communities: A Prospective Study of Six Online Cancer Forums. *Int J Comput Models Algorithms Med* 2011 Dec;2(2):1-22 [FREE Full text] [doi: [10.4018/jcmam.2011040101](#)] [Medline: [22022636](#)]
29. Wang X, Zhao K, Street N. Social Support and User Engagement in Online Health Communities. In: *Smart Health.*: Springer International Publishing; 2014 Presented at: International Conference for Smart Health; July 10-11, 2014; Beijing, China p. 97-110.
30. Borgatti S. Centrality and network flow. *Social Networks* 2005;27(1):55-71.
31. Freeman L. Centrality in social networks conceptual clarification. *Social Networks* 1978 Jan;1(3):215-239. [doi: [10.1016/0378-8733\(78\)90021-7](#)]
32. Myeni S, Fujimoto K, Cobb N, Cohen T. Content-Driven Analysis of an Online Community for Smoking Cessation: Integration of Qualitative Techniques, Automated Text Analysis, and Affiliation Networks. *Am J Pub Health* 2015;105(6):1206-1212. [doi: [10.2105/AJPH.2014.302464](#)]
33. MacQueen J. Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.*: University of California Press; 1967 Presented at: The Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1967; Berkeley, California.
34. Borgatti SP. Identifying sets of key players in a social network. *Comput Math Organiz Theor* 2006 Apr;12(1):21-34. [doi: [10.1007/s10588-006-7084-x](#)]
35. Kleinberg JM. Authoritative sources in a hyperlinked environment. *J ACM* 1999;46(5):604-632. [doi: [10.1145/324133.324140](#)]
36. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. In: *Computer Networks and ISDN Systems*. 1998 Apr Presented at: Seventh International Conference on the World Wide Web; April 14-18, 1998; Brisbane, Australia p. 107-117. [doi: [10.1016/S0169-7552\(98\)00110-X](#)]
37. Tausczik YR, Pennebaker JW. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *J Lang Soc Psych* 2009 Dec 08;29(1):24-54. [doi: [10.1177/0261927X09351676](#)]
38. Wang Y, Kraut R, Levine J. Eliciting and receiving online support: using computer-aided content analysis to examine the dynamics of online social support. *J Med Internet Res* 2015;17(4):e99 [FREE Full text] [doi: [10.2196/jmir.3558](#)] [Medline: [25896033](#)]
39. Nunes S, Ribeiro C, David G. The impact of time in link-based Web ranking. *Information Research* 2013;18(3):paper 586.

40. Yu P, Li X, Liu B. On the temporal dimensions of search. In: Proceedings of the 13th International World Wide Web Conference. 2004 Presented at: 13th International World Wide Web Conference; 2004; New York, NY p. 448-449. [doi: [10.1145/1013367.1013519](https://doi.org/10.1145/1013367.1013519)]

Abbreviations

HITS: hyperlink-induced topic search

IRRs: influential responding replies

OHC: online health community

Edited by G Eysenbach; submitted 17.05.15; peer-reviewed by S Myneni, M Ekstedt, E Børøsund, R Jones, K Zhao; comments to author 09.07.15; revised version received 26.08.15; accepted 07.10.15; published 01.12.15

Please cite as:

Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM

From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities

J Med Internet Res 2015;17(12):e271

URL: <http://www.jmir.org/2015/12/e271/>

doi: [10.2196/jmir.4705](https://doi.org/10.2196/jmir.4705)

PMID: [26627369](https://pubmed.ncbi.nlm.nih.gov/26627369/)

©Bradley Carron-Arthur, Kathina Ali, John Alastair Cunningham, Kathleen Margaret Griffiths. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 01.12.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

Chapter 3

Foreword

A limitation of the existing literature on matters central to this thesis was highlighted by the systematic review in Chapter 2. Studies that identify participation inequality in MHISGs have arbitrarily grouped users into three categories according to posting frequency (super users, contributors and lurkers, operationalised as the top 1%, intermediary 9% and bottom 90% of users respectively) based on a rule of thumb [13] that has long been applied to citizen engagement with social media. These studies did not acknowledge that the underlying distribution of user posts conforms to an inverse power-law. To address this gap, the third chapter of this thesis comprises a short paper reporting on an empirical study that replicated the previous research and formally investigated the question:

RQ 2: What is the mathematical distribution of engagement as measured by posting frequency among active users in an MHISG?

and contextualised the investigation within the relevant body of research including key literature on Zipfian distributions.

Throughout the remaining empirical studies in this thesis, users have been analysed in either high-, moderate- or low-engaged groups defined by cut-offs in the spectrum of engagement. This approach has been adopted for pragmatic, primarily statistical reasons, and to facilitate comparisons to previous research.



Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law



Bradley Carron-Arthur^{a,*}, John A. Cunningham^{a,b}, Kathleen M. Griffiths^a

^a National Institute for Mental Health Research, Australian National University, 63 Eggleston Road, Acton, Canberra, ACT 0200, Australia

^b Centre for Addiction and Mental Health, 33 Russell Street, T526 Toronto, ON, Canada

ARTICLE INFO

Article history:

Received 18 July 2014

Received in revised form 11 September 2014

Accepted 11 September 2014

Available online 28 September 2014

Keywords:

eHealth

Internet support group

Social network

Zipf's Law

90-9-1 principle

1% rule

ABSTRACT

Sustainable online peer-to-peer support groups require engaged members. A metric commonly used to identify these members is the number of posts they have made. The 90-9-1 principle has been proposed as a 'rule of thumb' for classifying members using this metric with a recent study demonstrating the applicability of the principal to digital health social networks.

Using data from a depression Internet support group, the current study sought to replicate this finding and to investigate in more detail the model of best fit for classifying participant contributions.

Our findings replicate previous results and also find the fit of a power curve (Zipf distribution) to account for 98.6% of the variance.

The Zipf distribution provides a more nuanced image of the data and may have practical application in assessing the 'coherence' of the sample.

© 2014 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/3.0/>).

1. Introduction

Online peer-to-peer support has many potential health benefits (Ziebland and Wyke, 2012). To date, systematic reviews have failed to find consistent evidence for the efficacy of online peer-to-peer support groups on health outcomes (Eysenbach et al., 2004; Griffiths et al., 2009). However, there is evidence that consumers value these groups (Horrigan et al., 2001) and there is increasing interest in identifying the key components of sustainable thriving online support groups (Young, 2013). It is generally agreed that one key component is highly engaged core members who contribute substantially to the community (Young, 2013). There is no consensus on what metrics should be employed to classify the contributions of members. Four studies have sought to identify highly engaged members in online peer-to-peer support groups using different combinations of metrics. These metrics include the number of posts made by members (Cobb et al., 2010; Jones et al., 2011; van Mierlo et al., 2012; van Mierlo, 2014), the number of threads initiated (Jones et al., 2011; van Mierlo et al., 2012), the number of different threads in which a member participates (Jones et al., 2011; van Mierlo et al., 2012), the level of connectedness to other members in the forum (Cobb et al., 2010) and time spent logged in (Jones et al., 2011). One metric common to them all was number of posts.

Recent research has used number of posts as a sole means of classifying members in Digital Health Social Networks (DHSN) with a peer-to-peer support group component (van Mierlo, 2014). The study investigated the 90-9-1 principle or the 1% rule. This rule describes a commonly reported phenomenon whereby the majority of content in an Internet community is produced by only 1% of the participants (referred to as 'superusers'), a minority of the content is produced by a further 9% of participants ('contributors') and 90% of people observe the content in the Internet community without actively participating ('lurkers') (Nielsen, 2014). The study sectioned the content attributed to these three groups and found that the sections contained 74.7%, 24.0% and 1.3% of the total posts in the DHSN respectively. It was concluded that the 90-9-1 principle applied to DHSN.

The DHSN study sought to verify the 90-9-1 principle rather than to determine the distribution which best fitted the data. Thus, the 90-9-1 principle may not provide the greatest accuracy in classifying participants in a DHSN. The aim of the current study is to further investigate the model of best fit for classifying participants in a DHSN, including but not limited to the 90-9-1 principle.

2. Method

This study used data from the peer-to-peer Internet support group – BlueBoard (blueboard.anu.edu.au). BlueBoard is predominantly used for peer-to-peer discussion about Depression (38.8% of content). It also includes forums on Bipolar Disorder (18.4%), Generalised Anxiety Disorder (5.0%), general discussion (22.1%) and other topics (15.7%).

* Corresponding author. Tel.: +61 02 6125 6825.

E-mail addresses: Bradley.Carron-Arthur@anu.edu.au (B. Carron-Arthur), John.Cunningham@anu.edu.au (J.A. Cunningham), Kathy.Griffiths@anu.edu.au (K.M. Griffiths).



The image shows the BlueBoard homepage. At the top left is the BlueBoard logo. To its right is a navigation menu with links for 'Emergencies Help', 'Register', 'FAQ', 'Calendar', and 'Today's Posts'. Below the menu is a search bar and a login section with fields for 'User Name', 'Password', and a 'Remember Me?' checkbox, along with a 'Log in' button. A central banner contains a welcome message and a warning about privacy. The main content area is a list of forum sub-forums, each with a folder icon and a link to 'Taking care of ourselves'.

BlueBoard

Emergencies Help Register FAQ Calendar Today's Posts Search

User Name Password Remember Me?

Welcome to BlueBoard!

BlueBoard is an online community for people suffering from depression or anxiety, their friends and carers, and for those who are concerned that they may have depression or anxiety and want some support. We hope that this bulletin board will enable people to reach out and both offer and receive help. The main thing we want you to know is that you are not alone! In order to post messages you will first need to [register](#) with a made-up alias.

IMPORTANT Please don't use a real name as part of your username or for privacy reasons we will have to disable your account. We just hate having to do this but protecting members' privacy is a really important aspect of BlueBoard. :) [Click here](#) to read more about usernames.

Forum

- BlueBoard Notices**
 - Sub-Forums: [Rules and Consents](#), [BlueBoard Notices](#), [Emergencies Help](#), [New members: What happened to Fred Smith?](#)
- Depression**
 - Sub-Forums: [Living with depression](#), [Taking care of ourselves](#)
- Bipolar Disorder**
 - Sub-Forums: [Living with bipolar disorder](#), [Taking care of ourselves](#)
- Generalised Anxiety**
 - Sub-Forums: [Living with generalised anxiety](#), [Taking care of ourselves](#)
- Social Anxiety**
 - Sub-Forums: [Living with social anxiety](#), [Taking care of ourselves](#)
- Panic Disorder**
 - Sub-Forums: [Living with panic disorder](#), [Taking care of ourselves](#)
- Obsessive Compulsive Disorder**
 - Sub-Forums: [Living with OCD](#), [Taking care of ourselves](#)
- Borderline Personality and Related Disorders**
 - Sub-Forums: [Living with borderline personality disorder](#), [Taking care of ourselves](#)
- Eating Disorders**
 - Sub-Forums: [Living with an eating disorder](#), [Taking care of ourselves](#)
- Caring for someone with a mental health problem**
 - Sub-Forums: [General](#), [Depression and Bipolar Disorder](#), [Anxiety Disorders](#), [Other disorders](#)
- General**
 - Sub-Forums: [Click chat](#), [Having a laugh](#), [Creative corner](#), [Suggestion Box](#)

Fig. 1. BlueBoard homepage.

BlueBoard is moderated by a team of paid personnel. Members are consumers and carers. BlueBoard’s homepage is shown in Fig. 1. The data used in this study included all posts generated between 1st October 2008 and the 23rd May 2014 (n = 131,004 by 2932 members). Posts made by moderators (n = 352 by 10 moderators) were not included in the analysis. Data collection procedures were approved by the Australian National University Human Research Ethics Committee.

In order to replicate the analysis conducted by van Mierlo (2014), we separately calculated the total number of posts made by the 1% of registered members who contributed the most, the next 9% and the final 90%. To investigate alternative models of fit for the data we graphed on a log–log scatterplot the total number of posts of each member ranked in order of those who made most to least posts and fitted a power curve using Microsoft Excel.

3. Results

The percentages of posts made by participants in each of the three Sections 1, 9, and 90 were 85.8%, 11.2% and 3.0% of the total number of posts respectively. The corresponding number of members in each section and the range in the number of posts made by members in that section are shown in Table 1.

A log–log scatterplot showing the frequency of posts made by each member ranked in descending order is presented in Fig. 2. The best fitting curve was found to have the function $f(x) = 63935x^{-1.427}$ with correlation coefficient $r = 0.993$ and a coefficient of determination of 0.986. This indicates that the model accounts for 98.6% of the variance.

4. Discussion

The current analysis broadly replicated the findings of van Mierlo (2014), that the top 1% of registered members contribute the vast majority of posts, the next 9% a minority and the last 90% very few. Thus, the 90-9-1 principle appears to provide a reliable means of broadly categorising participant contributions in a DHSN. However, the graph in Fig. 2 and the associated best fitting power curve provide an alternate and more precise means of describing the distribution. In fact, the distribution in Fig. 2 adheres to Zipf’s law – that the frequency of posts made by a member is inversely proportional to their rank in frequency. This is a widely observed phenomenon spanning areas such as linguistics, populations, income and internet traffic (Newman, 2006; Adamic and Huberman, 2002). This model gives a more nuanced image of the distribution. It shows a gradual reduction in contributions rather than a quantum leap at the boundary between superusers and contributors as the 90-9-1 principle implies. Researchers, developers and other stakeholders seeking to optimise the network effects associated with members who generate the highest levels of traffic in an Internet support group (van Mierlo, 2014) may benefit from the understanding that there is a predictable diminishing return associated with each individual member as opposed to categorical differences in types of users.

A range of explanations has been proposed to explain the occurrence of Zipfian distributions including, for example, the principle of least effort (Ferrer i Cancho and Sole, 2003), proportional growth processes (Gabaix, 1999) or a simple stochastic process (Miller et al., 1958). There is no consensus on which is correct and none allow a meaningful interpretation of the current data. However, a phenomenon associated

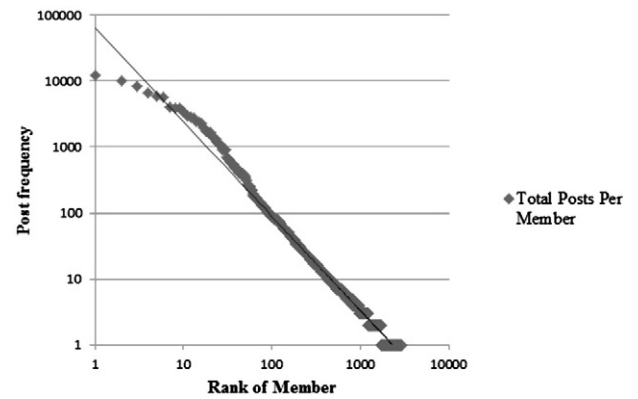


Fig. 2. Log–log scatterplot of the total posts made by each member ranked in descending order and a power curve which best fits the data.

with data which better fits the Zipfian distribution is that of greater ‘coherence’ in the sample (Cristelli et al., 2012). For example, ranking cities by population size in the USA fits the Zipfian distribution better than the European Union (EU). Furthermore, each individual country of the EU fits the distribution well in comparison to the EU as a whole, and conversely each individual state in the USA does not fit the distribution well in comparison to the USA as a whole. This is thought to reflect the time each has had to organically evolve as a collective unit (Cristelli et al., 2012). For Internet support groups, describing the distribution of engagement using the Zipfian distribution may allow researchers and developers to assess the coherence of the group versus the coherence of its subsets, such as the different forums within the group. In the current study, the best fit was found for the support group as a whole as opposed to any individual forum by topic.

Frequency of posts is one way of identifying highly engaged members in a network. It is not necessarily the most suitable method. Borgatti (2006) argues that key members in a network are most appropriately identified using the combination of metrics that identifies members whose engagement contributes the kind of value that reflects the reason they are being sought. In addition to the metrics which have been used in past research, future research may investigate other metrics such as the average word count of posts, time of day, regularity of posting or combinations of these. Since quantity does not necessarily reflect quality, content analysis of posts is required to determine if the highly engaged users are contributing informative and supportive content (Salem et al., 1997).

5. Conclusion

The 90-9-1 principle and Zipf’s Law both provide a means of describing the distribution in engagement of members by post frequency in the internet support group but Zipf’s law provides a more precise description of the data.

Acknowledgements

The authors would like to thank Anthony Bennett, Julia Reynolds and Kylie Bennett for their contributions to establishing and maintaining BlueBoard. We thank Anthony Bennett and Kylie Bennett for assistance in downloading the data. We also thank Professor David Hawking for his expert input regarding Zipf’s law. BlueBoard is funded by the Australian Government Department of Health. B. Carron-Arthur is supported by an Australian Postgraduate Award. K.M. Griffiths is supported by the Australian National Health and Medical Research Council (NHMRC) Research Fellowship 1059620.

Table 1
Posts and members in each section.

Percentile	Members (N)	Percentage of posts (N)	Range in the number of posts (N)
1 (1%)	1–74 (74)	85.8% (112,373)	11,994–142 (11,852)
2–10 (9%)	75–743 (669)	11.2% (14,669)	141–5 (136)
11–100 (90%)	744–7434 (6691)	3.0% (3,962)	5–0 (5)

References

- Adamic, L.A., Huberman, B.A., 2002. Zipf's Law and the Internet. *Glottometrics* 3, 143–150.
- Borgatti, S.P., 2006. Identifying sets of key players in a social network. *Comput. Math. Organ. Theory* 12, 21–34.
- Cobb, K.N., Graham, A.L., Abrams, D.B., 2010. Social network structure of a large online community for smoking cessation. *Am. J. Public Health* 100, 1282–1289.
- Cristelli, M., Batty, M., Pietronero, L., 2012. There is more than a power law in Zipf. *Sci. Rep.* 2.
- Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., Stern, A., 2004. Health related virtual communities and electronic support groups systematic review of the effects of online peer to peer interactions. *BMJ* 328.
- Ferrer I Cancho, R., Sole, R.V., 2003. Least effort and the origins of scaling in human language. *Proc. Natl. Acad. Sci. U. S. A.* 100, 788–791.
- Gabaix, X., 1999. Zipf's law for cities an explanation. *Q. J. Econ.* 114, 739–767.
- Griffiths, K.M., Calear, A.L., Banfield, M., 2009. Systematic review on Internet Support Groups (ISGs) and depression (1): do ISGs reduce depressive symptoms? *J. Med. Internet Res.* 11, e40.
- Horrigan, J.B., Rainie, L., Fox, S., 2001. *Online Communities: Networks That Nurture Long-Distance Relationships and Local Ties*. Pew Internet & American Life Project, Washington, DC.
- Jones, R., Sharkey, S., Smithson, J., Ford, T., Emmens, T., Hewis, E., Sheaves, B., Owens, C., 2011. Using metrics to describe the participative stances of members within discussion forums. *J. Med. Internet Res.* 13, e3.
- Miller, G.A., Newman, E.B., Friedman, E.A., 1958. Length-frequency statistics for written English. *Inf. Control.* 1, 370–389.
- Newman, M.E.J., 2006. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* 46, 323–351.
- Nielsen, J., 2014. Participation Inequality: Lurkers vs Contributors in Internet Communities, (URL: <http://www.nngroup.com/articles/participation-inequality/>. Accessed: 2014-06-05. (Archived by WebCite® at <http://www.webcitation.org/6Q7EwEncA>).
- Salem, D., Bogat, A., Reid, C., 1997. Mutual help goes on-line. *J. Community Psychol.* 25, 189–207.
- Van Mierlo, T., 2014. The 1% rule in four digital health social networks: an observational study. *J. Med. Internet Res.* 16, e33.
- Van Mierlo, T., Voci, S., Lee, S., Fournier, R., Selby, P., 2012. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter. *net. J. Med. Internet Res.* 14, e66.
- Young, C., 2013. Community management that works: how to build and sustain a thriving online health community. *J. Med. Internet Res.* 15, e119.
- Ziebland, S., Wyke, S., 2012. Health and illness in a connected world how might sharing experiences on the Internet affect people's health. *Milbank Q.* 90, 219–249.

Chapter 4

Foreword

This chapter presents the second of five empirical studies included in this thesis and the first of four that investigated an aspect of participation in BlueBoard followed by a comparative analysis of differentially-engaged users.

This study focused on BlueBoard users' personal characteristics including gender, age, location and status as either a consumer, carer or other. It was undertaken to address a gap in the literature identified in a systematic review of studies on depression ISGs [4] that had not been addressed by subsequent studies of these communities. The review identified that selection effects associated with surveys of ISG users may have biased the results of existing studies of the characteristics of depression ISG users and that none had used information collected during registration to analyse user characteristics more reliably.

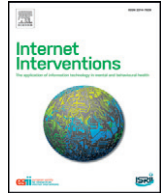
Accordingly, the study in this chapter sought to answer the questions:

RQ 3a: What are the characteristics of BlueBoard users?

RQ 3b: Are there systematic differences in the characteristics of users with different levels of engagement?

using information collected during registration. The comparative analysis of differentially-engaged users involved primarily a comparison between users who had contributed a single post and users of all other levels of engagement. The remaining three studies of this thesis focused largely on the higher-engaged category of users in their comparative analyses. The current study includes a survival analysis which I undertook of the retention data. The analysis suggests that the user characteristics employed in this study have only minor predictive value for user engagement.

The publication of the study in this chapter was markedly delayed for journal-related technical reasons unrelated to its content. The study itself was undertaken at the same time as the network modularity study in Chapter 5 which also incorporates user characteristics and prior to Chapters 6 and 7.



User characteristics and usage of an open access moderated internet support group for depression and other mental disorders: A prospective study



Kathleen M Griffiths^{a,*}, Bradley Carron-Arthur^b, Julia Reynolds^b, Kylie Bennett^a, Anthony Bennett^a

^a Research School of Psychology, College of Medicine, Biology and Environment, The Australian National University, Canberra, ACT 2601, Australia

^b Centre for Mental Health Research, College of Medicine, Biology and Environment, The Australian National University, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:

Received 16 July 2015

Received in revised form 11 March 2016

Accepted 25 November 2016

Available online 30 November 2016

Keywords:

Internet

Support group

Mental health

Depression

Anxiety

Virtual community

ABSTRACT

Background: Internet support groups (ISGs) for mental ill-health are common but little is known about the characteristics of users, the usage and predictors of ISG usage and if and how these change over time.

Aim: This study evaluated the attributes of a publically accessible ISG for depression and other mental disorders including: (1) the demographic and other characteristics of its users; (2) their patterns of usage; and (3) the factors which predict posts to and retention on the ISG.

Method: User characteristics (gender, age, user type, country and location of residence) were collected at the time of registration on the ISG BlueBoard (blueboard.anu.edu.au). All board log data were downloaded for the period October 2008 to May 2014. Predictors of post frequency and retention on the board were examined using logistic regressions. Other data were analysed using descriptive statistics.

Results: 2932 users contributed 131,004 posts to the ISG. The majority were female, aged 20 to 34 years, and mental health consumers. Although most users were city dwellers, 19% resided in rural or remote regions. Frequency of posts and retention on the board varied across users, with a moderate association between retention and number of posts. Growth in posts substantially exceeded the growth in new users over the monitoring period. Multivariate analysis demonstrated that consumers posted more often and remained longer than carers or others, and that younger users posted less often; however, the model predicted very little of the variance.

Conclusions: A small minority of active users are sufficient to ensure the sustainability and growth of an online mental health ISG. Further research is required to understand why so many support group members limit their contributions to one or a very small number of posts and what factors predict and promote active engagement and long-term retention in virtual mental health communities.

© 2016 The Author(s). 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/>).

1. Introduction

Internet support groups (ISGs) are accessible and popular (Dosani et al., 2014; Fox and Duggan, 2013; Parikh and Huniewicz, 2015) and have the potential to provide valued social support to individuals with depression and other common mental disorders (Barak et al., 2008; Griffiths et al., 2015). While there is uncertainty regarding the effectiveness and safety of ISGs (Eysenbach et al., 2004; Griffiths et al., 2009a; Hoybye et al., 2010; Parikh and Huniewicz, 2015; Rice et al., 2014; Takahashi et al., 2009), recent high quality evidence suggests that such support groups may improve mental health outcomes (e.g., (Griffiths et al., 2012, Ali et al., 2015)) and increase users' sense of empowerment (Crisp et al., 2014), self-esteem (Crisp et al., 2014)

and perceived quality of life (Crisp et al., 2014). However, there is little systematically collected evidence on the characteristics of those who use ISGs for depression or mental ill-health, or what determines the level of participation and retention of users in these groups (Griffiths et al., 2009b).

There is some evidence that the predominant users of depression support groups are consumers (Houston et al., 2002; Powell et al., 2003; Salem et al., 1997; Alexander et al., 2003; Nimrod, 2012) who are primarily in their mid-20s to 40s (Dosani et al., 2014; Houston et al., 2002; Nimrod, 2012; Powell et al., 2003); there is mixed evidence regarding gender of users (Dosani et al., 2014; Fekete, 2002; Houston et al., 2002; Nimrod, 2012; Powell et al., 2003; Salem et al., 1997; Takahashi et al., 2009). However as noted by Griffiths et al. (2009b), a limitation of most previous studies of public depression support groups is that they have been derived from surveys posted on ISGs or by inferring the users' status from posts. Thus, the data collected from these studies is typically either restricted to those ISG members who remain on the board and who chose to participate in the surveys, or conclusions

* Corresponding author at: Research School of Psychology, College of Medicine, Biology and Environment, The Australian National University, 2601, Australia.

E-mail address: Kathy.Griffiths@anu.edu.au (K.M. Griffiths).

are based on inferences of unknown validity. A more valid approach would analyse data collected at the time of registration.

To date, few studies have investigated patterns of usage of open access mental health support groups across forum topics or over time (Griffiths et al., 2009b) based on all registered users and registration details. Although one group has undertaken a study of membership duration and its predictors in a cross-sectional survey (Nimrod, 2012), as noted above there are limitations to the conclusions that can be drawn from such methodologies. We are not aware of any systematic studies of the retention patterns for all users of an online mental health support group. Nor – with the exception of cross-sectional studies – are we aware of studies that have systematically investigated individual differences in mental health support group usage by ISG members or the predictors of any such differences based on details at registration. We have recently reported the distribution of posts across an entire online support group for depression and related disorders (blueboard.anu.edu.au), finding that they conformed to a Zipfian distribution (Carron-Arthur et al., 2014). However, the study did not investigate individual differences in detail.

Accordingly, the current study sought to document: (1) the demographic and other characteristics of users of the publically accessible bulletin board on registration; (2) the patterns of usage on the board; and (3) the factors which predict usage of the board and retention on the board.

2. Method

Data were collected from the database of the peer-to-peer ISG BlueBoard. Ethics clearance was obtained from the ANU Human Research Ethics Committee prior to the establishment of the board to enable the investigators to undertake research investigating the characteristics of the board and its users.

2.1. The Internet support group: BlueBoard (blueboard.anu.edu.au)

This service was provided by the National Institute for Mental Health Research/Centre for Mental Health Research at the Australian National University with funding from Australia's Department of Health. BlueBoard comprised 10 forums including: (1) Eight *condition* forums each focused on a different mental disorder (depression, bipolar disorder, generalised anxiety, social anxiety, panic disorder, obsessive compulsive disorder, borderline personality disorders and eating disorders). Each of the condition forums comprised two sub-forums: "Living with [condition, e.g., depression]" and "Taking care of our ourselves"; (2) A *carer* forum ("Caring for someone with a mental health problem") comprising four sub-forums: "General", "Depression and Bipolar disorder"; "Anxiety disorders" and "Other disorders"; and (3) a *general* forum comprising four sub-forums 'Chit-chat', "Having a laugh", "Creative corner" and "Suggestions box". The Board was moderated by consumers who were trained and supervised by a registered clinical psychologist (JR). The Board was originally established as a mood disorder support group in 2003 but was closed in 2007 and 2008 due to lack of funding. The current study is focused on the second phase of the service. All forums were established on 1 October 2008 except the Obsessive Compulsive, Borderline Personality and Eating Disorder forums which were established on the 1 June 2009, 1 March 2010 and 30 July 2012 respectively. Further details of the Board can be found elsewhere (Griffiths et al., 2015).

2.2. Measures

User characteristics were collected at the time of registration on BlueBoard and included: age range, gender, country of residence, location of residence (rural/capital city/other city) and type of user (consumer/carers/other). Usage data were collected by downloading all

posts, their time stamps, and forum and sub-forum details, for the period 1 October 2008 and 23 May 2014.

2.3. Analyses

Analyses were undertaken using SPSS Statistics Version 22.0. User characteristics and usage (number of posts contributed, and user retention on the board in months) were analysed using descriptive statistics. Simple bivariate relationships between user characteristics and usage were assessed using Kruskal-Wallis and Mann-Whitney tests. Multivariate analysis of predictors of usage were examined using logistic regression analyses. Retention data were further analysed using Cox proportional hazards regressions with (i) no censoring; (ii) right censoring of users who posted in the last 2 weeks of the data collection period and (iii) right censoring of users who posted in the last 12 weeks.

3. Results

A total 4823 individuals registered on BlueBoard (excluding those who were banned for spamming or related activity) in the relevant period, of whom 2932 contributed at least one post. The focus of this paper is on those registrants who contributed one or more posts; they will be referred to here as 'users'.

3.1. Characteristics of users

Table 1 summarises the characteristics of the BlueBoard users who provided demographic and other data. Missing data ranged from 8.6% (gender) to 15% (participant type). Two-thirds of users were women and the majority were consumers. Although the Board was used by older people including some aged over 75 years, the majority of users were aged between 20 and 34 years. Users resided in 76 different countries, with most living in Australia. The next most frequent user locations were the United States and the United Kingdom. Approximately half of all users lived in a capital city but a substantial minority (almost 19%) resided in rural or remote areas.

Table 1

Characteristics of users (≥ 1 post, $n = 2932$) at registration. Values are numbers (percentages) unless otherwise stated.

Characteristic	n (%)	
Gender ($n = 2680$)		
Women:	1812 (67.6%)	Chi-square (1) = 1812, $p < 0.001$
Age category ($n = 2662$)		
<20 yrs	174 (6.5%)	
20–34 yrs	1489 (55.9%)	
35–49 yrs	737 (27.7%)	
50–64 yrs	243 (9.1%)	Chi-square (4) = 2691.07, $p < 0.001$
65 yrs +	19 (0.7%)	
Location/rurality ($n = 2602$)		
Rural/remote	483 (18.6%)	
Capital city	1367 (52.5%)	Chi-square (2) = 473.50, $p < 0.001$
Other city	752 (28.9%)	
Country ($n = 2671$)		
Australia	2195 (82.2%)	
United States	202 (7.6%)	
United Kingdom	70 (2.6%)	
Canada	29 (1.1%)	
New Zealand	16 (0.6%)	
India	15 (0.6%)	
Other	144 (5.4%)	
User type ($n = 2493$)		
Consumer	1664 (66.7%)	
Carer	238 (9.6%)	
Other	591 (23.7%)	Chi-square (2) = 1327.48, $p < 0.001$

3.2. Patterns of usage

3.2.1. Distribution of posts across forums and sub-forums

Users contributed a total of 131,004 posts during the 68-month study period. Of these, the majority were posted in the condition forums ($n = 89,667$; 68.4%) and a substantial minority were in the general forum ($n = 38,780$, 29.6%); the remainder were posted to the carer forum ($n = 2557$, 2%).

Table 2 provides further details of the distribution of posts across forums and sub-forums. Of all the posts to condition forums, the overwhelming majority were in the “Living with...” sub-forums. Of the eight condition forums, the majority of posts were made to the depression forum followed by the bipolar forum. Similarly, within the carer forum, the majority of posts were made to the depression and bipolar disorder sub-forums. In addition, a significant minority of the carer forum posts were made to the carer general sub-forum. Finally, the majority of the posts in the general forum were made to the chit-chat sub-forum (‘general off-topic chit chat’), although a significant minority of the general forum posts were attributable to ‘creative corner’ where participants shared ‘poetry, short stories or articles...or talk about other creative endeavours’.

3.2.2. Distribution of posts and new users across time.

Total posts increased from 900 in 2009 to 60,251 in 2013, a growth factor of 67 times (6700%). This reflected the high growth rates in posts in most forums (see Table 3). During the same period the number of new users increased by from 207 to 622, a growth of 3 times (300%). Cumulative users grew from 252 at the end of 2009 to 2637 at the end of 2013, a growth factor of 10.5. Since not all of the early users remained active in 2013, this is an overestimate of users for 2013. Despite this, the growth factors in posts far exceeded the growth in registered users.

3.3. Individual differences in usage: posts

We have previously reported that the distribution of posts for all registered members conformed to a Zipfian (power) distribution (Carron-Arthur et al., 2014) with total posts for individuals ranging from 0 to 11,994. The current study focused on the participants who made at least 1 post. Of these 10% of users ($n = 239$) contributed 93.8% of the posts, each contributing between 19 and 11,994 posts. They will be referred to here as ‘active users’. A total of 41.2% of users ($n = 1207$) contributed one post each only; they will be designated here as ‘one-off users’. The remaining group of users ($n = 1432$, 48.8% of users) posted between 2 and 18 posts and will be referred to as ‘multiple users’.

3.3.1. Interrelationship between posts across forums

The intercorrelations between post frequencies for different forums were in most cases statistically significant but very small. However, as noted above, over 40% of the users posted only once and almost half of all users posted fewer than 19 times, limiting the likelihood that posts would be distributed across different forums. Accordingly, the data were examined for the 349 ‘active users’ who posted at least 19 times. For these users, there was a statistically significant moderate association between the posts in the condition forum (combining all sub-forums) and the combined posts in the general forum (including the Chit Chat, Creative Corner, Having a laugh, and Suggestion Box sub-forums) (Spearman rho = 0.51, $p < 0.001$). There was also a small significant correlation between the frequency of posts for the condition and carer forum posts (Spearman rho = 0.36, $p < 0.001$). Similarly, there was a significant, moderate relationship between the posts in the carer and general forums (Spearman rho = 0.57, $p < 0.001$). Thus, the more an active user posted in one forum, the greater their number of posts in other forums.

3.3.2. Association between user characteristics and posts.

Given the markedly skewed distribution of posts, bivariate analyses of the effect of different users characteristics on usage were undertaken using non-parametric tests. These analyses found no difference in post frequency for male and female users (Mann-Whitney U = 754.187, $p = 0.074$). However, there was a significant difference in total posts across different user types ($n = 2493$, Kruskal Wallis Chi-square(2) = 27.43, $p < 0.001$). Pairwise comparisons, adjusted for multiple comparisons, showed that consumers posted more frequently than carers (adjusted $p < 0.001$) or ‘others’ (adjusted $p < 0.001$). There was no significant difference in post frequency for carers and others (adjusted $p = 1$). There was a significant difference in usage across the five age groups (Kruskal Wallis: Chi-square (4) = 10.92, $p = 0.027$) with younger users (<20 years) posting less frequently than those aged 35 to 49 years ($p = 0.019$ after adjustment). Uncorrected comparisons suggested that those aged <20 years posted less frequently than users aged between 20 and 34 years and between 50 and 65 years (uncorrected $p = 0.018$) but these differences were no longer statistically significant following correction for multiple comparisons ($p = 0.18$). Finally there was no statistical difference in total posts based on whether the participant resided in a rural location, a capital city or another city (Kruskal-Wallis, Chi-square (2) = 3.90; $p = 0.14$).

Normal probability plots (P-plots) of the residuals indicated that the assumption of normality was not met either for the total posts or log-transformed total posts. It was therefore not appropriate to undertake a linear regression analysis. Predictors of post frequency were then investigated using the three-category classification (active, multiple and one-off users) described above. However, when an ordinal regression

Table 2
Number and percentage of posts within forums and sub-forums (Oct 2008–May 2014).

Forum type	Forums	n (%)	Sub-forum	n (%)
Condition forums $n = 89,667$	Depression	52,904 (59.0%)	Living with...	86,235 (96.2%)
	Bipolar disorder	24,829 (27.7%)	Taking care of ourselves	3432 (3.8%)
	Generalised anxiety disorder	6898 (7.7%)		
	Social anxiety disorder	2144 (2.4%)		
	Borderline personality disorder	1363 (1.5%)		
	Eating disorder	807 (0.9%)		
	Panic disorder	507 (0.6%)		
	Obsessive compulsive disorder	215 (0.2%)		
Carer forum $n = 2557$			Depression & bipolar	1536 (60.1%)
			General	719 (28.1%)
			Anxiety	244 (9.5%)
			Other	58 (2.3%)
			Chit-chat	29,016 (74.8%)
General forum $n = 38,780$			Creative corner	7948 (20.5%)
			Having a laugh	1087 (2.8)
			Suggestion box	729 (1.9%)

Table 3
Posts and new users as a function of year.

	2008 (Oct–Dec)	2009	2010	2011	2012	2013	2014 (Jan–May)
Depression	188	529	1948	4292	15,787	25,187	4973
Bipolar disorder	33	59	662	1281	9071	11,628	2095
Generalised anxiety disorder	15	64	84	548	2655	2770	762
Social anxiety	17	30	58	329	1062	558	90
Panic disorder	0	14	15	127	98	203	50
OCD ^a	N/A	1	8	25	90	90	1
Borderline personality disorder ^a	N/A	N/A	32	229	343	660	99
Eating disorder ^a	N/A	N/A	N/A	N/A	512	142	153
Caring for someone with a mental health problem	4	44	117	346	1356	575	115
General (Chit chat, Creative corner, Having a laugh, Suggestion Box)	128	159	414	855	16,893	18,438	1893
Total posts each year	385	900	3338	8032	47,867	60,251	10,231
Total new users each year	45	207	453	576	734	622	295

^a The OCD, borderline personality and eating disorder forums were commenced on 1 June 2009, 1 March 2010 and 30 July 2012.

was undertaken with gender, age, location and user as independent variables, the assumption of parallel lines was violated and there were a substantial number of cells with zero or small sample sizes primarily in the 'Active' category. Similarly, a multinomial regression was not appropriate due to the number of zero and small sample size cells.

Accordingly the usage data was dichotomised into two categories (1 = one post only users; 2 = multiple post users) and a logistic regression undertaken entering gender, age, location, and user status as independent variables, the latter three being coded as dummy variables (see Table 4 (i)). This analysis demonstrated an overall significant effect for the model (Chi-square(9) = 19.85, $p = 0.02$), with a higher frequency of posts for consumers compared to carers and others, fewer posts for those aged <20 years compared to the 20 to 34 and 35 to 50 year old groups, and a trend towards lower posts among rural and remote residents compared with their city counterparts ($p = 0.051$). Gender was not a predictor of multiple compared to one-off posts. The Hosmer-Lemeshow test indicated that the model provide an adequate fit to the data (Chi-square(8) = 7.79, $p = 0.35$). However, the Nagelkerke R^2 value was only 0.011, indicating that the effect was very weak.

Dummy variables: Age reference group ≤ 20 yrs; Location reference group = Capital city; User type reference group = Consumer.

3.4. Individual differences in usage: retention on the board

Retention of users on the board, calculated as the time between registration and the last user activity at the time of data download, ranged from 0 to 5.5 years. The distribution was highly negatively skewed with half of the users engaged with the board for <1 day (49.0%; $n = 1438$), 30.3% visiting for between 1 day and a month ($n = 887$), 17% for between one month and 1 year ($n = 498$), and the remaining 3.7% ($n = 109$) returning to the Board for more than one year. These four categories of users will be denoted: minimal, short-term, medium-term

Table 4
Predictors of post frequency: logistic regression for a two-category dependent variable (1 post vs multiple posts).

Independent variables	Unstandardized coefficient		Standardized coefficient Exp beta	Wald	Sig
	B	SE			
Gender	0.09	0.09	1.10	1.03	0.311
Age 20–34 yrs	–0.41	0.17	0.66	5.62	0.02
Age 35–49 yrs	–0.44	0.18	0.65	5.85	0.02
Age 50–64 yrs	–0.31	0.21	0.73	2.13	0.14
Age 65+ yrs	–0.55	0.51	0.58	1.20	0.27
Other city	–0.01	0.10	1.01	0.003	0.196
Rural remote	–0.22	0.12	0.80	3.80	0.051
Carer	–0.29	0.14	0.75	4.21	0.04
Other	0.20	0.10	1.22	4.05	0.04

and long-term users. Overall, the correlation between posts and retention was 0.68 (Spearman rho) and 0.53 (Kendall tau).

3.4.1. Association between user characteristics and retention

Since the distribution of user retention periods was skewed, bivariate analyses of the effect of different user characteristics on retention were undertaken using non-parametric tests. The results of these analyses were broadly consistent with those for user posts, although retention was longer for female than male users ($n = 2680$, Mann-Whitney $U = 741,027$, $p = 0.015$). Retention differed across user types ($n = 2493$, Kruskal Wallis Chi-square(2) = 52.41, $p < 0.001$) with Bonferonni adjusted pairwise comparisons demonstrating that consumers remained active on the board longer than carers (adjusted $p < 0.001$) or 'others' user types (adjusted $p < 0.001$). There was no significant difference in retention for carers and others after adjustment for multiple comparisons (adjusted $p = 0.084$). Retention differed across location of user (Kruskal Wallis Chi-square(2) = 6.66, $p = 0.036$), with adjusted pairwise comparisons demonstrating that users from rural/remote areas remained active for longer than those from non-capital cities (adjusted $p = 0.033$). A similar pattern was seen for rural/remote users compared to those from capital cities, but the effect was no longer statistically significant after Bonferonni adjustment (adjusted $p = 0.13$). There was a significant difference in retention across age groups categorized into 5 groups ($n = 2662$, Kruskal Wallis Chi-square(2) = 13.06, $p = 0.011$). Without adjustment for multiple comparisons those aged <20 years remained active on the board for less time than those aged 35–49 years ($p = 0.044$) or 50–65 years ($p = 0.009$); in addition those aged 20–34 remained active less time than those aged 35 to 49 years ($p = 0.009$). However, none of these effects remained statistically significant after Bonferonni adjustment.

Normal probability plots (P-plots) of the residuals indicated that the assumption of normality was not met either for the retention data or the log-transformed retention data, the latter being distributed bimodally. An attempt was made to investigate predictors of retention by undertaking a multinomial analysis on the four-category classification of the retention periods described above (namely minimal, short-term, medium-term and long-term users). To reduce the number of cells with zero sample sizes for this variable, before undertaking the analysis each of the independent variables were dichotomised based on the pattern of differences demonstrated by the bivariate analyses (Age: 1 ≤ 35 years 2 = 35 years and above; User: 1 = consumer 2 = carer or other; Location: 1 = capital or other city; 2 = rural/remote) before undertaking the analysis. However, the assumption of parallel lines was violated.

The retention data were therefore dichotomised into two categories (1 = active for 1 day or less ($n = 1438$); 2 = active for >1 day ($n = 1492$)) and a logistic regression undertaken using gender, age, location, and user status as independent variables. The latter three factors were coded as dummy variables (see Table 5 (i)). The analysis demonstrated

Table 5

Predictors of retention: logistic regression for a two-category dependent variable (retention <1 day vs 1 or more days).

Independent variables	Unstandardized coefficient		Standardized coefficient Exp beta	Wald	Sig
	B	SE			
Gender	0.05	0.09	1.05	0.33	0.57
Age 20–34 yrs	−0.26	0.17	0.77	2.19	0.14
Age 35–50 yrs	−0.32	0.18	1.37	3.01	0.08
Age 50–64 yrs	−0.15	0.21	0.86	0.52	0.47
Age 65+ yrs	0.52	0.52	1.68	0.98	0.32
Other city	−0.05	0.10	1.05	0.23	0.63
Rural remote	−0.17	0.11	0.85	2.28	0.13
Carer	0.60	0.14	0.55	17.34	<0.001
Other	0.36	0.10	1.43	13.09	<0.001

Dummy variables: Age reference group ≤20 years yrs; Location reference group = Capital city; User type reference group = Consumer.

an overall significant effect for the model (Chi-square(9) = 35.43, $p < 0.001$), and greater retention by consumers compared to carers ($p < 0.001$) and other users ($p < 0.001$). There were no other significant predictors. The Hosmer-Lemeshow test indicated that the model provided an adequate fit to the data (Chi-square(7) = 1.76, $p = 0.97$). However, the Nagelkerke R^2 value was only 0.02, indicating that the effect was very weak. A similar pattern of findings was demonstrated when the retention data were dichotomised into retention for <1 month and retention of 1 month and greater (Chi-square(9) = 43.54, $p < 0.001$, Hosmer-Lemeshow Chi-square(8) = 6.27, $p = 0.62$, Nagelkerke $R^2 = 0.037$) with a consumers more likely than carers and others to use the Board for at least a month (consumer vs carer Wald (1) = 9.75, $p = 0.002$; consumer vs other Wald (1) = 21.78, $p < 0.001$). As a final consistency check three survival analyses (Cox proportional hazards regression with (i) no censoring, right censoring for last (ii) 2 and (iii) 12 weeks of data collection) were undertaken on the data. In each case the findings were consistent with those of the logistic regression except that the trend in the logistic regression for those aged 20 years to remain for less time than users aged 35 to 50 years achieved statistical significance in the survival analyses.

4. Discussion

Using details collected at registration we found that over the almost 5.75 year period of the study, a total of 2932 people contributed a total 131,004 posts to a publicly accessible, moderated online support group for depression and related disorders. The majority of users were women, aged between 20 and 34 years, and consumers (with direct experience of mental ill-health). Most resided in a city and lived in Australia, although visitors were from 76 different countries and 19% of users were from rural or remote regions. Posts were primarily to condition forums (and the 'Living with Depression' forum in particular), although a substantial minority of posts were concerned with more general topics (in particular 'Chit-Chat'). There was a very large growth in posts over the monitoring period that substantially exceeded the growth in new users. Of the active BlueBoard users, those who posted more in one forum also posted more in other forums. Retention periods on the board varied across users with a moderate association between retention and number of posts. Bivariate analyses suggested that consumers posted more often and remained more engaged than other users and that younger users (<20 years) posted less actively than older users and tended to remain engaged with the board for less time. Further, these uncontrolled analyses indicated that women remained on the Board over a longer period than men, as did people from rural and remote areas compared with their city counterparts - although the latter effects were not statistically significant after adjustment for multiple comparisons. Multivariate analyses confirmed consumers posted more often and engaged with the board for longer,

and that young people posted less often and engaged over a shorter period than some older age groups. However gender and location effects on retention were not sustained after controlling for other factors; rural and remote users showed a trend to lower post frequency.

The distribution of characteristics among users of BlueBoard was consistent with previous reports that consumers (Houston et al., 2002; Powell et al., 2003; Salem et al., 1997; Alexander et al., 2003) comprise a greater proportion of users of online depression support groups than their counterparts. Significantly, in contrast to most previous research the methodology of the current study used registration data that did not rely on potentially unrepresentative surveys or unvalidated inferences based on the content of posts. Together with the findings in the current study that consumers posted more frequently and over a longer period of time, the findings from the current and previous studies suggest that online support groups for depression and related disorders are more important for those with lived personal experience of the condition than for carers. Alternatively, it is possible that the contributions from carers initially fell short of the critical mass required to sustain carer peer engagement thereby compromising further growth of carer involvement, that the strong consumer presence reduced carer sense of ownership and identification with the Board and that carers might prefer a space dedicated to their needs and priorities.

Our finding that there were more users among the younger (20–34 years) than the older cohorts is consistent with the findings of previous survey studies (Dosani et al., 2014; Houston et al., 2002; Powell et al., 2003). Could this pattern simply reflect the age distribution of the population? According to Australian Bureau of Statistics census figures, 22% of Australians were aged between 20 and 34 years in 2012–13 (Australian Bureau of Statistics, 2015a). Clearly then this age group was substantially overrepresented among users of the board which accounted for 55% of the current sample. Further, this overrepresentation cannot be explained by the relatively small disparity in Internet access among age groups (Australian Bureau of Statistics, 2014). The higher prevalence of common mental disorders among the younger cohort may have contributed at least in part to this pattern (Australian Bureau of Statistics, 2008). However, given the magnitude of the observed difference in frequency of users in each of these age groups it is likely that age exerted an independent effect on initial engagement with the online support group that was not explained by other factors such as internet access and mental health status.

Young users <20 years were more likely than those aged 20 to 50 years to make only one post and posted on the board over a shorter period than those aged 35 to 50 years. This group is apparently less engaged by the activities on the board. There may be a number of explanations for this. The age group comprised only a small proportion of users. A lack of availability on the board of age-matched peers with shared life experience and interests may have served as a barrier to frequent and sustained participation. Age can confer a perception of authority, mitigating against the operation of equal peer relationships. Alternatively, mirroring face-to-face barriers, young people may have difficulty in or be reluctant to articulate their emotional problems. On the other hand, peer support groups may in fact be less engaging for young people than for older consumers. Further research is required to investigate the usage and drivers of online peer-to-peer support among this age group.

Whereas previous studies have reported mixed findings with respect to the gender of users (Dosani et al., 2014; Fekete, 2002; Houston et al., 2002; Powell et al., 2003; Salem et al., 1997; Takahashi et al., 2009), the current study showed a clear predominance of women (67%) among users of the online support group. These figures may reflect the higher rates of depression in women, with studies consistently reporting a prevalence of depression in women that is twice that of men (Kessler, 2003). The lack of gender effects on post frequency and retention on the board after controlling for other factors suggests that once they post on an online support group men and women engage to a similar extent.

The finding that a substantial minority of users of the bulletin board resided in rural or remote regions is important. Overall, census figures indicate that almost two-thirds (66%) of Australians resided in a capital city in June 2014 whereas only a little over one-half (52%) (Australian Bureau of Statistics, 2015b) of BlueBoard users were from a capital city. It is unclear if this higher prevalence of users from outside the main cities reflects a preference among these citizens for self-help or to provide help to others, a lower access to face-to-face mutual support groups, the lack of accessible professional services in rural areas, or the impact of some other factor. It does however suggest that an online service of this type may be an acceptable and feasible source of support for citizens residing in rural areas who otherwise lack either face-to-face peer support or accessible professional services.

Although the majority of users were from Australia, the findings emphasise the potential reach of an online support group such as this with users residing in 76 different countries. If, as has been proposed, the concept of universality of illness is an important factor in recovery (Yalom and Leszcz, 2005), then it is conceivable that the commonality of experiences shared by people from diverse backgrounds might contribute to recovery among users of the service.

It was notable that the majority of posts were to the condition forums within which there was less emphasis on 'Taking care of ourselves' forums. The "Living with" forums are for "reaching out to others and sharing your experiences. You are not alone with [the condition]" whereas the 'Taking care of ourselves' forum is about "What do you do to take care of yourself? Here is a supportive forum where you can share the things you do to feel good." The difference in popularity of the forums may reflect a focus by users not only on sharing the impacts of the condition and seeking support but also on providing support to those in distress rather than focusing specifically on communicating positive strategies." Moreover, it may be that some of the posts in the "Living with" condition include sharing positive strategies as part of the support provided by users to others. Recent research on a cancer support group has found that members were less likely to respond to messages expressing positive emotion (Lewallen et al., 2014). Thus, the response patterns of members may encourage users to focus on negative content. On the other hand, the relatively high rate of posts to the general forum comprising 'Chit-Chat' indicates that users were not solely focused on the negative effects of their condition. For active users there was a positive relationship between posts to the condition and general forums suggesting that these users were forming connections that transcended their disorders. It is possible that such interactions have a positive effect on mental health through the pursuit of safe but normalising interactions at a time of the user's choosing. The latter may be difficult to sustain in the course of face-to-face interactions, the timing and duration of which are less easily controlled and the nature of which may be less predictable and less supportive.

The large growth in posts relative to a smaller growth in new users suggests that an increase in posts may generate greater engagement among existing users. Recent research indicates that receiving online problem- and emotion-focused support increases a person's willingness to reciprocate by providing support (Lin et al., 2015). It has previously been suggested that insufficient activity may compromise the sustainability of support groups (Jones et al., 2004). Conversely, however, it is thought that too much activity might discourage users by creating a less personal space and information load among visitors (Jones et al., 2004). Clearly, the current support group was still in the growth phase after 5 years.

In the current sample, only 3.7% of users engaged with the bulletin board for more than one year. It is instructive to compare this figure with the distribution of users reported in the cross-sectional survey undertaken by Nimrod et al. (Nimrod, 2012) in which 40% of respondents had been members for over one year. This raises questions about the use of survey methodology to explore the characteristics of online support groups. More particularly it casts doubt on Nimrod et al.'s conclusion that participants in their survey were 'quite representative of members

of online depression support groups' ((Nimrod, 2012), p. 1255). Nevertheless, both the latter study and the current study reported higher retention among older participants and neither reported a gender effect on retention.

4.1. Limitations and future research

Although there was a consistent effect of user type on post frequency and duration of usage, the effect was extremely weak. This suggests that other factors not identified in the current study are the major contributors to differences in usage. A limited range of personal characteristics were measured at registration. Future research should incorporate additional variables at registration (e.g., symptoms severity) to enable a more comprehensive investigation of the predictors of usage and to identify the factors which predict usage on online support groups.

The current study did not examine the relationship between user characteristics or usage and outcomes, either positive or negative. Even if ISGs are associated with an positive health outcomes at a group level, or as reported previously that higher usage is linked to better outcomes (Houston et al., 2002), it cannot be assumed that the effects apply to all individuals. A substantial minority of users in a research-specific ISG reported "feeling upset that they could not help the other members of the ISG more" (29%) and 'frustration that they could not meet the other members of the ISG in person' (17%) (Crisp and Griffiths, 2016). Other studies of ISGs have reported the potential for 'a downward depressive spiral triggered by aggravated psychological burden' (Takahashi et al., 2009) and one study reported an association between online forum use and increased suicidal ideation among young people (Dunlop et al., 2011). Further, although an analysis of user posts to BlueBoard are characterized by a predominantly positive user perspective, potential disadvantages of participation occur when an individual does not feel that they belong in the group due to the severity of their problems or other personal characteristics such as age (Griffiths et al., 2015). Accordingly, there is a need not only to undertake further investigations of ISG user characteristics, usage and outcomes individually but also to examine how they interact and in particular how a user's characteristics and fit within the composition of the support group affect the outcomes of participation.

5. Conclusion

Online support groups attract consumers with depression and related disorders including young adults and those from rural and remote areas. Although only a small minority of users were active, the evidence from the current study suggests that a small group of regular users is sufficient to ensure the sustainability and growth of the group. Further research is required to understand why so many support group members limit their contributions to one or a very small number of posts, what factors predict and promote active engagement and long-term retention in virtual mental health communities and which if any user characteristics and usage factors affect health outcomes.

Authors' contributions

KG conceived the study, undertook the statistical analyses and wrote the paper. BC pre-processed the data, undertook the survival analyses, and contributed to the interpretation of the data. JR, AB and KB contributed to the data collection. All authors edited the paper and have read and approved the final manuscript.

Declaration of interest

KG established BlueBoard and KG, JR KB and AB were responsible for the provision of the BlueBoard service. None of the authors derived personal financial benefit from the operation of the service.

Acknowledgements

At the time of the study Kathleen Griffiths was supported by NHMRC Fellowship 1059620. Bradley Carron-Arthur was supported by an Australian Postgraduate Award. BlueBoard was supported by funding from the Australian Department of Health.

References

- Alexander, S.C., Peterson, J.L., Hollingshead, A.B., 2003. Help is at your keyboard: support groups on the internet. In: FREY, L.R. (Ed.), *Group Communication in Context: Studies of Bona Fide Groups*, second ed. Lawrence Erlbaum Associates, London.
- Ali, K., Farrer, L., Gulliver, A., Griffiths, K.M., 2015. Online peer-to-peer support for young people with mental health problems: a systematic review. *JMIR Ment. Health* 2, e19.
- Australian Bureau of Statistics, 2008. *National Survey of Mental Health and Wellbeing: Summary of Results*. Australian Bureau of Statistics, Canberra.
- Australian Bureau of Statistics, 2014. *Household use of Information Technology, Australia, 2012–13*. 8146.0. ABS, Canberra.
- Australian Bureau of Statistics, 2015a. *Estimated resident population ERP by Region, Age & Sex, 2001–2013*. [Online]. ABS, Canberra (Available: <http://stat.abs.gov.au/> [Accessed [Last accessed 4 June 2015.]
- Australian Bureau of Statistics, 2015b. *Regional Population Growth, Australia, 2013–14*. Report no 3218.0. ABS, Canberra.
- Barak, A., Boniel-Nissim, M., Suler, J., 2008. Fostering empowerment in online support groups. *Comput. Hum. Behav.* 24, 1867–1883.
- Carron-Arthur, B., Cunningham, J.A., Griffiths, K.M., 2014. Describing the distribution of engagement in an internet support group by post frequency: a comparison of the 90-9-1 principle and Zip's law. *Internet Interv.* 1, 165–168.
- Crisp, D.A., Griffiths, K.M., 2016. Reducing depression through an online intervention: benefits from a user perspective. *JMIR Ment. Health* 3.
- Crisp, D.A., Griffiths, K.M., Mackinnon, A.J., Bennett, K., Christensen, H., 2014. An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. *Psychiatry Res.* 216, 60–66.
- Dosani, S., Harding, C., Wilson, S., 2014. Online groups and patient forums. *Curr. Psychiatry Rep.* 16, 507.
- Dunlop, S.M., More, E., Romer, D., 2011. Where do youth learn about suicides on the internet, and what influence does this have on suicidal ideation? *J. Child Psychol. Psychiatry* 52, 1073–1080.
- Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., Stern, A., 2004. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Br. Med. J.* 328.
- Fekete, S., 2002. The internet — a new source of data on suicide, depression and anxiety: a preliminary study. *Arch. Suicide Res.* 6, 351–361.
- Fox, S., Duggan, M., 2013. *Health Online 2013*. PewResearchCentre, Washington, D.C.
- Griffiths, K.M., Calear, A.L., Banfield, M., 2009a. Systematic review on internet support groups (ISGs) and depression (1): do ISGs reduce depressive symptoms? *J. Med. Internet Res.* 11.
- Griffiths, M.K., Calear, L.A., Banfield, M., Tam, A., 2009b. Systematic review on internet support groups (ISGs) and depression (2): what is known about depression ISGs? *J. Med. Internet Res.* 11, e41.
- Griffiths, K.M., Mackinnon, A.J., Crisp, D.A., Christensen, H., Bennett, K., Farrer, L., 2012. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. *PLoS One* 7, e53244.
- Griffiths, K.M., Reynolds, J., Vassallo, S., 2015. An online, moderated peer-to-peer support bulletin board for depression: user-perceived advantages and disadvantages. *JMIR Ment. Health* 2, e14.
- Houston, T.K., Cooper, L.A., Ford, D.E., 2002. Internet support groups for depression: A1-year prospective cohort study. *Am. J. Psychiatry* 159, 2062–2068.
- Hoybye, M.T., Dalton, S.O., Deltour, I., Bidstrup, P.E., Frederiksen, K., Johansen, C., 2010. Effect of internet peer-support groups on psychosocial adjustment to cancer: a randomised study. *Br. J. Cancer* 102, 1348–1354.
- Jones, Q., Ravid, G., Rafaeli, S., 2004. Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration. *Inf. Syst. Res.* 15, 194–210.
- Kessler, R.C., 2003. Epidemiology of women and depression. *J. Affect. Disord.* 74, 5–13.
- Lewallen, A.C., Owen, J.E., Bantum, E.O.C., Stanton, A.L., 2014. How language affects peer responsiveness in an online cancer support group: implications for treatment design and facilitation. *Psycho-Oncology* 23, 766–772.
- Lin, T.-C., Hsu, J.S.-C., Cheng, H.-L., Chiu, C.-M., 2015. Exploring the relationship between receiving and offering online social support: a dual social support model. *Inf. Manag.* 52, 371–383.
- Nimrod, G., 2012. The membership life cycle in online support groups. *Int. J. Commun.* 6, 23.
- Parikh, S.V., Huniewicz, P., 2015. E-health: an overview of the uses of the internet, social media, apps, and websites for mood disorders. *Curr. Opin. Psychiatry* 28, 13–17.
- Powell, J., McCarthy, N., Eysenbach, G., 2003. Cross-sectional survey of users of internet depression communities. *BMC Psychiatry* 3, 19.
- Rice, S.M., Goodall, J., Hetrick, S.E., Parker, A.G., Gilbertson, T., Amminger, G.P., Davey, C.G., McGorry, P.D., Gleeson, J., Alvarez-Jimenez, M., 2014. Online and social networking interventions for the treatment of depression in young people: a systematic review. *J. Med. Internet Res.* 16.
- Salem, D.A., Bogat, G.A., Reid, C., 1997. Mutual help goes on-line. *J. Commun. Psychol.* 25, 189–207.
- Takahashi, Y., Uchida, C., Miyaki, K., Sakai, M., Shimbo, T., Nakayama, T., 2009. Potential benefits and harms of a peer support social network service on the internet for people with depressive tendencies: qualitative content analysis and social network analysis. *J. Med. Internet Res.* 11.
- Yalom, I.D., Leszcz, M., 2005. *The Theory and Practice of Group Psychotherapy*. Basic Books, NY.

Chapter 5

Foreword

This chapter presents the third of five empirical studies included in this thesis and the second of four that investigated an aspect of participation in BlueBoard followed by a comparative analysis of differentially-engaged users.

This study focused on BlueBoard community structure through a social network analysis. It sought to answer the questions:

RQ 4a: Are there sub-groups (modules) within the BlueBoard community structure that comprise users who communicate with each other relatively more frequently than with other users?

RQ 4b: Do users within the same sub-group share a higher degree of common characteristics, including posting frequency in disorder-specific forums, than would be expected by chance?

RQ 4c: Within each sub-group and with respect to the previously ascertained common characteristic, are there systematic differences in that characteristic between higher- and lower-engaged users?

In this study, all users in the network were divided into sub-groups (modules) based on a modularity algorithm that optimises the density of connections (communications between users) within sub-groups relative to between sub-groups. In order to extract any commonalities and differences within and between the modules, the user characteristics employed in Chapter 4 were entered as potential predictors of sub-group allocation in a multinomial logistic regression, together with user registration date and posting frequency in each of the disorder-specific forums.

The findings of this study suggest that high-engaged users are central to the network and potentially drive the observed cohort-like structure given the timing of their registration (significantly earlier than the median in each module). The fact that this structure does not mirror the pre-existing structure of BlueBoard's disorder-specific sub-forums informed the

design of subsequent studies in this thesis and in particular highlighted the need to avoid study designs that impose a structure on the data a priori, for example, by using a pre-formulated coding scheme for analysing content. Consistent with the findings reported in Chapter 4, user characteristics were again found not to be a strong predictor of differences in participatory patterns. Thus, user characteristics were not used as a predictor in any subsequent studies in this thesis, with the exception of the study in Chapter 7 which was restricted to consumers.

Note: The paper in this chapter makes reference to '[Personal communication by Kathleen M Griffiths, 2016]'. This refers to the paper included in Chapter 4 that had been delayed.

Original Paper

Community Structure of a Mental Health Internet Support Group: Modularity in User Thread Participation

Bradley Carron-Arthur¹, BPsych (Hons); Julia Reynolds¹, B.A. (Hons), MPsych (Clin); Kylie Bennett¹, BA (Hons), BSc; Anthony Bennett¹, BAppSc; John Alastair Cunningham^{1,2}, PhD; Kathleen Margaret Griffiths¹, PhD

¹National Institute for Mental Health Research, Research School of Population Health, The Australian National University, Acton, Australia

²Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:

Bradley Carron-Arthur, BPsych (Hons)
National Institute for Mental Health Research
Research School of Population Health
The Australian National University
63 Eggleston road
Acton, 2601
Australia
Phone: 61 2 6125 6825
Fax: 61 2 6125 0733
Email: Bradley.Carron-Arthur@anu.edu.au

Abstract

Background: Little is known about the community structure of mental health Internet support groups, quantitatively. A greater understanding of the factors, which lead to user interaction, is needed to explain the design information of these services and future research concerning their utility.

Objective: A study was conducted to determine the characteristics of users associated with the subgroup community structure of an Internet support group for mental health issues.

Methods: A social network analysis of the Internet support group BlueBoard (blueboard.anu.edu.au) was performed to determine the modularity of the community using the Louvain method. Demographic characteristics age, gender, residential location, type of user (consumer, carer, or other), registration date, and posting frequency in subforums (depression, generalized anxiety, social anxiety, panic disorder, bipolar disorder, obsessive compulsive disorder, borderline personality disorder, eating disorders, carers, general (eg, “chit chat”), and suggestions box) of the BlueBoard users were assessed as potential predictors of the resulting subgroup structure.

Results: The analysis of modularity identified five main subgroups in the BlueBoard community. Registration date was found to be the largest contributor to the modularity outcome as observed by multinomial logistic regression. The addition of this variable to the final model containing all other factors improved its classification accuracy by 46.3%, that is, from 37.9% to 84.2%. Further investigation of this variable revealed that the most active and central users registered significantly earlier than the median registration time in each group.

Conclusions: The five subgroups resembled five generations of BlueBoard in distinct eras that transcended discussion about different mental health issues. This finding may be due to the activity of highly engaged and central users who communicate with many other users. Future research should seek to determine the generalizability of this finding and investigate the role that highly active and central users may play in the formation of this phenomenon.

(*JMIR Mental Health* 2016;3(2):e20) doi:[10.2196/mental.4961](https://doi.org/10.2196/mental.4961)

KEYWORDS

internet; support group; social network; modularity; mental health; super user

Introduction

Online peer-support is a popular source of health information and social support. Findings suggest that in a 1-year period, 18% of the Internet users in the USA sought information online that was provided by a peer [1] and 8% actively sought a response or provided support to another peer by engaging in an online health community [2]. Annually, 28% of Internet users have sought mental health specific information online [3]. Consequently, Mental Health Internet Support Groups (MHISGs) can comprise thousands of users who are actively participating to varying degrees [4,5]. MHISGs are popular and have high potential to play a role in the management of mental illness. Research on MHISGs must address various questions concerning the nature of MHISGs such as “Who uses them?” and “How are they used?” in order to fully benefit from this potential [6]. Recent research on the demographic characteristics of MHISG users has identified differences in prevalence, engagement, and retention of users with different characteristics, such as age, gender, location and consumer or carer status [Personal communication by Kathleen M Griffiths, 2016]. This information is important in understanding to whom do MHISGs have greater appeal. The willingness to engage and the outcome of participating in the MHISG may be different for each user depending on whom they interact with, however, it is also important to understand the social dynamics of how users engage with each other.

From a sociological perspective, the principle of homophily suggests that those who group together, in this instance by communicating most often with each other, tend to share common characteristics [7]. If the premise of peer-support is a shared experience, then it is plausible that homophily may be an important underlying factor in the community structure of the MHISG, that is, the community structure of the MHISG may comprise various subgroups, each consisting of users with higher proportions of shared characteristics than in other subgroups.

Many characteristics may affect the formation of subgroups in the MHISG, with some being more relevant than others. The most commonly observed factors influencing the people in interaction are age, gender, and location [3]. These factors are also influential across large-scale online social networks [4]. Specifically, in the domain of MHISGs, there are other factors, which may be important. Different mental health conditions are characterized by different symptoms and experiences [8]. From a psychological perspective, these are fundamental distinctions and they form the basis for different treatments. One might hypothesize that users in the MHISG with similar health concerns would seek to interact with each other, that is, people with depression concerns would provide peer-support to other people with depression, and not anxiety. However, people engaging in peer-support through MHISGs have the autonomy to interact with whomever they choose. These naturally occurring dynamics are currently unknown and a greater understanding of this area is needed. This information may empower community managers to take informed decisions concerning the design of MHISGs. Understanding these natural inclinations also provides a basis for future research to design

studies and form hypotheses about relevant factors, which if altered, may affect the outcome of participation and subsequently the potential utility of these communities.

To determine user grouping among the social network of the MHISG, it is recommended to conduct an analysis of its modularity [9]. Modularity is a measure that identifies subgroups in a social network by applying an algorithm designed to find a structure, which optimizes the number of communications within each module compared with the number of communications between different modules. The result of such an algorithm is the assignment of nodes (users) to modules (subgroups), which have a greater density of edges (communications) between them compared with nodes in other modules. It may be possible to use this algorithm in order to determine subgroups of users who engage in higher amounts of peer-support with each other than other users. Using these groups as an outcome, it may be possible to determine whether certain user characteristics are associated with those groupings. To the best of our knowledge, no study has yet investigated modularity in the MHISG. This study aims to determine the community structure of the MHISG through modularity and to explore the user characteristics associated with the resulting structure.

Methods

BlueBoard

The data used in this study were obtained from the publicly available Internet support group—BlueBoard (blueboard.anu.edu.au) established by the National Institute of Mental Health Research at the Australian National University. BlueBoard users must register and provide consent for their data to be used for research in order to participate in the MHISG. Peer-to-peer discussion on BlueBoard takes place anonymously via forum postings, which cover a range of topics, including depression (38.8% of posts), bipolar disorder (18.4%), generalized anxiety disorder (5.0%), chitchat and general discussion (22.1%), and other topics (15.7%). Posts dated between October 1, 2008, and May 23, 2014 were included in this study and were in a thread with posts given by two or more users ($n=130,582$ by 2652 users). BlueBoard is moderated by paid personnel who monitor content and enforce rules, for example, by editing posts to remove personally identifying information. BlueBoard moderators do not operate as facilitators of conversation, but post content occasionally regarding rules or other administrative matters. Moderator posts ($n=352$ by 10 moderators) were not included in the analysis. Data collection procedures were approved by The Australian National University Human Research Ethics Committee.

Measures. User characteristics included age (measured in 5 year brackets, eg, 25–29); gender (female, male); type of user (consumer, carer, other); location (capital city, other city, rural or remote region); registration date; and the number of posts in each of the subforums of BlueBoard (depression, generalized anxiety, social anxiety, panic disorder, bipolar disorder, obsessive compulsive disorder, borderline personality disorder, eating disorders, carers, general (eg, “chit chat”), and suggestions box).

Age, gender, location, and type of user were self-identified at the time of registration on BlueBoard. The last recorded activity of users was not more than 1 month or 1 year post registration for 86 and 97% of users respectively, thus suggesting that the data likely remained accurate for the majority of users throughout the period of the study. Data on age was grouped into three categories (<25, 25–34, >34) for the analysis to eliminate singularities in the Hessian matrix occurring in brackets above 60 years, with low counts. The three age categories contained approximately one-third of the users each. The term “consumer” refers to a person with depression, anxiety, or other mental health problems, and the term “carer” refers to a nonprofessional carer. The frequency of the posts in each of the subforums was tallied during the entire study duration, that is, from October 1, 2008 to May 23, 2014 for all subforums except the obsessive compulsive, borderline personality, and eating disorder forums, which were established on the June 1, 2009, March 1, 2010, and July 30, 2012, respectively.

Data Analysis

Modularity. Social network analysis was undertaken using Gephi 0.8.2. software [10]. Edges within nodes were defined as undirected communications between each user who had posted in the same thread. The modularity algorithm utilized was the widely-used Louvain method [11], which has the fastest computational time of any comparable algorithm and excellent performance in detecting communities [12]. The resolution was set to the default value 1.0 as this provided the highest modularity score (0.273).

Multinomial Logistic Regression. A multinomial logistic regression analysis was conducted to determine the user factors

that were significantly associated with the modularity outcome. There were 449 individuals who left at least one of the demographic questions unanswered while registering for BlueBoard. Little’s Missing Completely at Random test was not significant, indicating that the data was missing completely at random; accordingly, they were not included in the analysis.

Visualization. In order to explore the results patterns, graphs of the data underlying significant effects were created using pivot tables and charts in Microsoft Excel. To further explore the temporal factor associated with the registration date, a dynamic social network analysis was conducted. This required a timestamp to be associated with the creation of each edge in the social network. The time associated with the creation of each edge was the time a user first posted content in a thread. This edge was created only between the new user and users who had already posted in the thread. A visualization of the edges being created between nodes was generated using Gephi 0.8.2. software and TechSmith Jing screen recording tool [13].

Results

Modularity

The modularity algorithm produced 11 separate modules (See Table 1), out of which 6 modules contained less than 10 users. The latter modules were isolated from the giant component of the social network as they involved threads in which only new users not connected to the larger social network posted. The remaining 5 modules comprised between 328 and 954 users, which made 1977 and 67,590 posts, cumulatively. All subsequent analyses are concentrated on these five main modules as outcomes.

Table 1. Module sizes

Module	N (%)	Posts
1	434 (16.4)	1977
2	954 (36.0)	15,954
3	393 (14.8)	39,720
4	525 (19.8)	67,590
5	328 (12.4)	5300
6	8 (0.3)	27
7	2 (0.1)	2
8	2 (0.1)	2
9	2 (0.1)	5
10	2(0.1)	3
11	2 (0.1)	2
Total	2652 (100)	130,582

Multinomial Logistic Regression

Multinomial logistic regression is used to find the odds of being allocated to each of the different dependent variable outcomes based on a number of factors as predictors. In this analysis, the outcomes were the five different modules. One of the outcomes should be used as a reference category for comparison with the

other outcomes. In this case, we chose to use Module 4 because its users had contributed the highest number of posts. This decision was made before obtaining any knowledge regarding the number order, we labeled them with. In multinomial logistic regression each of the factors are used to predict the relative odds of persons from the reference group and the comparison group being allocated to each of the two groups as the predictive

factors change. In this analysis, several significant effects were found and each of the parameter estimates is shown in [Table 2](#). In this table, the odds ratios, which are the exponents of B, show the relative odds of being allocated to each outcome group as compared with the reference module (thus Module 4 is not included).

With respect to all independent variables in the analysis as predictors of the modularity outcomes, for each unit change, the odds of a person being allocated to the comparison group as opposed to the reference group changes by a factor of the odds ratio. As such, an odds ratio of < 1 indicates that as the

score of the predictor increases, the odds of a person being included in the comparison module decreases. An odds ratio > 1 indicates that as the score of the predictor increases, the odds of a person being included in the comparison module increases.

Overall, the final model fits the data significantly better than the null model (Chi square = 4146.4, $p < .001$). The classification accuracy of the model was 84.2% and the effect size was large (Nagelkerke $R^2 = 0.891$). The addition of one variable, registration date, improved the model classification accuracy by 46.3%. Without this variable, the effect size was much smaller (Nagelkerke $R^2 = 0.119$).

Table 2. Significant parameter estimates for the multinomial logistic regression of registration date, age, gender, location, type of user, and frequency of posts in the subforums on the dependent variable modularity.

Module ^a	Predictor	B	Standard error	Wald	p^b	Odds ratio
1	Registration date	-0.033	0.001	620.975	<.001	0.968
	Subforum: depression	-0.048	0.018	7.063	.008	0.953
	Subforum: carers	-0.782	0.280	7.834	.005	0.457
2	Registration date	-0.022	0.001	403.895	<.001	0.978
	Subforum: generalized anxiety disorder	-0.082	0.028	8.402	.004	0.921
	Subforum: borderline personality disorder	0.069	0.026	7.272	.007	1.071
	Subforum: suggestions	0.437	0.217	4.061	.044	1.548
3	Registration date	-0.009	0.001	171.921	<.001	0.991
	Subforum: suggestions	-0.165	0.075	4.750	.029	0.848
5	Registration date	0.008	0.001	149.971	<.001	1.008

^aModule 4 was used as the reference category

^bAll effects degrees of freedom = 1

Registration Date. There was a significant parameter estimate for the relationship between registration date and each of the module outcomes as shown in [Table 2](#). For comparing Modules 1–3 with the reference group, the odds ratios of registration date have values < 1 . This indicates that a person would be 0.968, 0.978, and 0.991 times as likely to be included in the groups 1–3, respectively, compared with the reference group (Module 4) for each day post registration. The opposite was true for Module 5 relative to the reference group. This indicates that a person would be 1.008 times more likely to be included in Module 5 than Module 4 for each day post registration.

User Characteristics. Across the three demographic variables and user type, there were no significant parameter estimates.

Frequency of Posting in Subforums. Based on the frequency of posts in the 11 different subforums, there were 6 significant parameter estimates across 5 different subforums. These are shown in [Table 2](#). For the comparison of Module 1 with the reference group, the odds ratios reveal that posting more in either the depression subforum or the carers subforum indicated that a person was more likely to be included in the reference group. For the comparison of Module 2 with the reference group, the odds ratios reveals that posting more in the generalized anxiety disorder subforum indicated that a person was more likely to be included in reference group. The opposite was true

for posting in the borderline personality disorder subforum and suggestions subforum. For the comparison of Module 3 with the reference group, posting more in the suggestions subforum indicated that a person was more likely to be included in the reference group. There were no significant parameter estimates for the frequency of posts in subforums in Module 5.

Visualization

Registration Date. The graph shown in [Figure 1](#) displays the number of users who registered with BlueBoard during each month from October 2008 to May 2014. Users are grouped by module. This graph supports the pattern of results found in the regression analysis. It shows that the five modules have five sequential time periods in which most of the users who signed up during that period were classified within that group. The distribution of new registrations in each of the four most recent subgroups loosely resembles a normal distribution.

A video showing the sequence in which edges were created between nodes is available as [Multimedia Appendix 1: BlueBoard social network growth time lapse](#). A graph representing this dynamic visualization is displayed in [Figure 2](#). Both show the progression of new communications occurring between users of BlueBoard, primarily between users of the same subgroup during each era progressing from 1 to 5.

Figure 1. Number of new users who registered each month between October 2008 and May 2014, grouped by module.

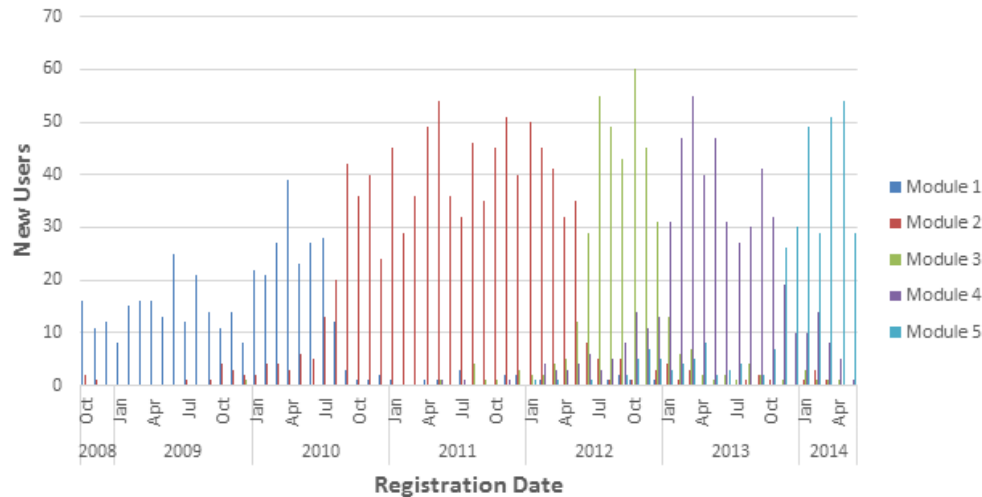
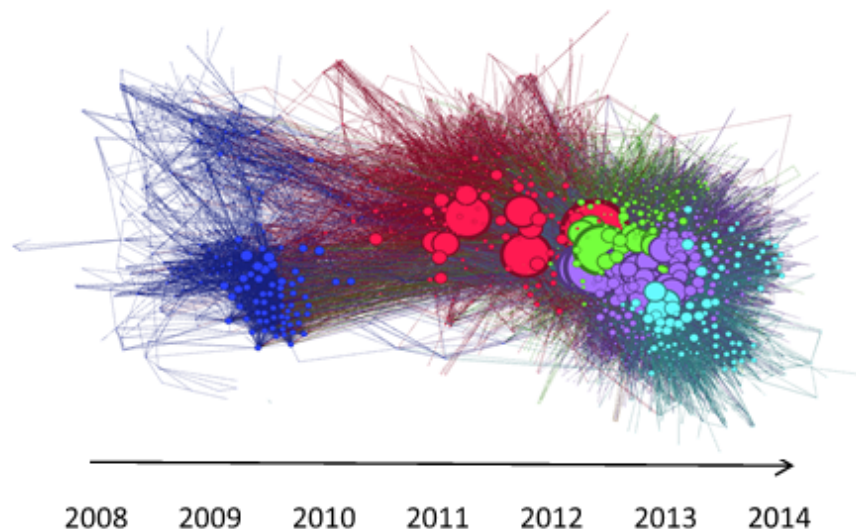


Figure 2. Graphical representation of the sequence of edges created between nodes. Each node is represented by a colored circle, and the nodes are colored according to their module. The size of each node corresponds to its degree (number of connections with other nodes). The layout was determined by the algorithm ForceAtlas 2 [15]. This algorithm places nodes, which have more edges between them, closer together. The arrow on the graph gives a general indication as to the order in which new edges are added to the network as time progresses.



Further Investigation of Registration Date

Since modularity was so strongly associated with registration date, we initiated further analyses to investigate the other factors associated with registration date that might explain the modularity pattern. Based on research, which suggests that online community development follows a life-cycle [14] and that certain “core users” play a vital role from the inception of that development [15,16], we hypothesized that there may be highly active and central users whose registration date is earlier than the majority of other users in each module. For this, we tested whether the top 10 users in each module, ranked by (1) total post frequency and (2) eigenvector centrality (a measure of network centrality, which identifies the most influential nodes [17]), registered significantly earlier than the median registration

date for each module. The results of these analyses are presented in Table 3. For total frequency of posts, we found that on average the top ranked users registered significantly earlier than the median registration time in all five modules ($\alpha < .05$). The case for eigenvector centrality was similar, except for the first module. This occurred despite the fact that, across BlueBoard as a whole, there was no significant difference between the average registration date of the top 10 users and the median for either total post frequency ($p=.40$) or eigenvector centrality ($p=.39$). In addition, there was no correlation between total post frequency and registration date (Spearman rho = 0.01, $p=.60$). Contrary to the pattern in each module, there was a significant positive correlation between registration date and eigenvector centrality (Spearman rho = 0.37, $p<.001$).

Table 3. T-tests of the mean difference (days) between the median registration date in each module and the average registration date of the top 10 users ranked by (1) total post frequency and (2) eigenvector centrality.

Module	Total post frequency			Eigenvector centrality		
	Mean difference (days)	<i>t</i>	<i>p</i> ^a	Mean difference (days)	<i>T</i>	<i>p</i> ^a
1	272	-3.56	.006	140	-1.15	.28
2	317	-3.34	.009	293	-2.74	.023
3	145	-3.32	.009	123	-2.64	.027
4	347	-5.96	<.001	377	-8.56	<.001
5	55	-2.84	.019	137	-2.85	.019

^aAll effects degrees of freedom = 9

Discussion

This study constitutes the first social network analysis of a mental health Internet support group in which the community structure was determined quantitatively through analysis of modularity. We investigated whether several user characteristics were associated with the resulting modularity outcome. The findings of this analysis provide a new perspective on how users engage in peer-support in MHISGs.

Principal Findings

We found that the community structure of the Internet support group BlueBoard comprised five main modules. Although there were several statistically significant parameter estimates across the different factors for this outcome, registration date contributed the most to the predictive power of the model. Statistically and visually, this factor stood out in the results. The pattern of results suggests that BlueBoard has progressed through a series of generations or eras. There were some minor differences in these generations in the degree to which their users posted in different subforums, but these frequencies did not differ substantially from the overall frequencies for BlueBoard reported elsewhere [Personal communication by Kathleen M Griffiths, 2016].

These results shed light on the nature of peer-support in MHISGs. They suggest that people who join the MHISG may communicate most with those who register around the same time. While this is not surprising, an important finding is the fact that registration date takes precedence over other factors such as demographic characteristics and type of mental health issue in predicting group membership in the MHISG. It raises the possibility that the social interactions of MHISGs are not largely affected by these characteristics. However, it is too early to draw a definitive conclusion as other factors may underpin the observed results.

In order to interpret the findings of this study, we considered whether artificial factors may have impelled the observed progression through each of the five subgroups. We considered two salient factors—external advertising and internal structural changes. Advertising of BlueBoard has occurred mainly via links from online mental health information hubs such as MindHealthConnect.org.au and bluepages.anu.edu.au. Following BlueBoard's establishment, postcard flyers were soon mailed to general practitioners to be displayed in waiting rooms.

Subsequently, there has been a gradual increase in the number of user registrations on BlueBoard. Therefore, recruitment did not appear to be a probable explanation. Further, with respect to internal sources, there were three subforums (obsessive compulsive, borderline personality, and eating disorders) that were introduced at different stages after BlueBoard's establishment. As there has been little uptake of these forums and they do not correlate with the progressions between the five subgroups, we did not consider this to be a probable explanation. We are not aware of any other developments or improvements that may have resulted in the observed findings. For this reason, we focused on the pattern of results involving the date of registration by highly engaged and central users in each module relative to the majority of other users. This pattern suggests that these users may have some role in the formation of this generation-like structure. However, further research is needed to test this hypothesis and to investigate if these findings generalize to other MHISGs.

Related Research

This study involved the first analysis of its kind for the MHISG. However, we are aware of a study involving an Internet support group for diabetes, which conducted a similar analysis [18]. This study sought to determine if a modularity analysis could be applied to an online health community and generate meaningful results by creating a formula, which was designed to measure the quality of the modularity outcome. This formula was based on the principle of homophily [7], such that greater similarity among the characteristics (eg, diagnosis) of users in each module resulted in a higher score. The study found that the modularity outcome was associated with the number of years since a user was diagnosed with the condition, indicating the time elapse since diagnosis was similar for users within each module. If as might be expected, the time a person takes to join an Internet support group after being diagnosed is relatively invariant; the findings of this study may have implications for our own. We did not measure time since diagnosis in our study. However, it is possible that the significant effect of registration date is confounded with and attributable to time since the diagnosis period. Alternatively, the diabetes study results can be explained by time of registration.

Our results suggested that type of health concern was not strongly linked to modularity outcome. By contrast, Chomutare et al.'s [18] formula produced a higher score for diagnosis of diabetes type rather than the time since diagnosis, indicating

that the former is the stronger determinant of the modularity outcome. This apparent difference in results might imply that the nature of peer-support in a mental health group is less strongly determined by specific health concerns than in a diabetes Internet support group. Alternatively, it could indicate that time since diagnosis has a much smaller effect on the modularity outcome as compared with the registration date in a diabetes Internet support group, or both.

The Role of Highly Active and Central Users in MHISGs

The observed pattern of highly active and central members registering early in each group in our study is consistent with research which suggests that these users play a vital role in the development of the community at an early stage [16]. The broader literature on online health communities report that “core users” engage in activities of building community by, for example, welcoming newcomers and communicating with many different people [16]. This finding was based on action research on the community #hcsma and was followed by a quantitative study of the same community, which suggested that core users could be identified as those who have the highest frequency of posts and network centrality [15]. A prospective study of a depression Internet support group suggests that these core users are veterans of the community who increasingly become “active help providers” after an initial period in which they are supported by the others in the community [19]. Thus, the findings from the current study interpret that each module represents an era in which several highly active users communicated with many other new users who registered at the same time regardless of whether they had similar characteristics (as measured in this study) or not, and that these core users played a key role in sustaining the community over time.

Limitations and Future Research

Although BlueBoard contains a range of subforums for different mental health topics, BlueBoard is predominantly used for discussion on depression. Thus, the generalizability of the current findings to other MHISGs is uncertain and in particular the modularity outcome may differ in MHISGs, which have an evenly spread distribution of posts across different mental health conditions. BlueBoard does not contain subforums for all types of mental health issues. Given the possibility that some forums are, therefore, not used for their intended purpose, the pattern of results may differ in MHISGs with a different variety of subforums. A more refined representation of the social network could also be achieved through collection of systematic data on directed communications between users.

The demographic characteristics of BlueBoard users [Personal communication by Kathleen M Griffiths, 2016] are similar to those of depression Internet support groups reported elsewhere [19]. However, the applicability of the current findings to MHISGs comprising members with markedly different demographic characteristics, such as those dedicated to young people, is unknown. MHISGs including medical professionals as moderators and or active participants might also be

characterized by markedly different social dynamics. Accordingly, further research focusing on a range of MHISG types is required to gain a greater understanding of the generalizability of the current findings. Future studies may benefit by modularity in MHISGs to collect and analyze a greater array of user characteristics including diagnosis, time since diagnosis, symptom severity, digital skills, and other characteristics that may reveal motivations of the “core users.”

The role of highly active and influential members is an important area for future research. There are multiple ways of measuring participation in an online health community including some specifying peer-leader roles [20]. We used broad measures in this study (posting frequency and eigenvector centrality), which may not capture the specific nature of different individuals’ contribution to the observed results. Future research with a more specific focus may consider other predefined peer-leadership roles such as “hubs” and “community builders,” who being high frequency posters, are also known for connecting many users and maintaining conversation, respectively [20]. In MHISGs where the identity of users is not anonymous, the role of users who act as hubs or bridges across multiple social networks should also be considered [21,22]. Concurrently, it is also important to understand which characteristics are associated with users who take up these roles. We recently conducted a study of BlueBoard to investigate the user characteristics associated with higher engagement than a single post [Personal communication by Kathleen M Griffiths, 2016], with consumers being found to be more highly engaged than carers. Further research is required to investigate the factors predicting the very highest levels of user engagement and other measures of peer-leadership in online health communities [20]. One previous study has compared the characteristics of the top 1% of users (“superusers”) ranked by posting frequency across two smoking cessation Internet support groups and found no differences between them [23]. A study with higher statistical power may be required to detect significant differences and common characteristics among such a small group of users. As super users are communicating with people who have a range of different mental health concerns, it is possible that super users have multiple or more complex diagnoses, which enable them to relate to and support the other bulk of users who have more specific issues or one-time needs for peer-support. Alternatively, they may have conditions such as bipolar disorder, which result in high activity levels with greater engagement in the community.

Conclusion

The community structure of the Internet support group BlueBoard comprised five main subgroups that occurred in sequence resembling generations of the MHISG. These groups were largely invariant in their demographic characteristics and the extent to which they communicated about different mental health issues. The community structure formation may be related to the contributions of the most active and central users who registered early as compared with other users in each group.

Acknowledgments

B. Carron-Arthur is supported by an Australian Postgraduate Award. K.M. Griffiths is supported by the Australian National Health and Medical Research Council (NHMRC) Research Fellowship. BlueBoard is supported by funding from the Australian Department of Health.

Conflicts of Interest

Kathleen Griffiths established BlueBoard; and Kathleen, Julia, Kylie, and Anthony are responsible for the provision of the BlueBoard service.

Multimedia Appendix 1

BlueBoard social network growth timelapse.

[[SWF File, 7MB - mental_v3i2e20_app1.swf](#)]

References

1. Fox S. PewResearchCenter. 2011 Feb 28. Peer-to-peer Health Care URL: <http://www.pewinternet.org/2011/02/28/peer-to-peer-health-care-2/> [accessed 2016-04-21] [WebCite Cache ID 6gvYqmuPt]
2. Fox S, Duggan M. : PewResearchCenter; 2013 Jan 15. Health Online 2013 URL: <http://www.pewinternet.org/2013/01/15/health-online-2013/> [accessed 2016-04-21] [WebCite Cache ID 6gvZEJUQr]
3. Fox S, Jones S. PewResearchCenter. 2011. The social life of health information URL: <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/> [accessed 2016-04-21] [WebCite Cache ID 6gvZTSOXw]
4. Carron-Arthur B, Cunningham J, Griffiths K. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions* 2014 Oct;1(4):165-168. [doi: [10.1016/j.invent.2014.09.003](https://doi.org/10.1016/j.invent.2014.09.003)]
5. van Mierlo T. The 1% rule in four digital health social networks: an observational study. *J Med Internet Res* 2014;16(2):e33 [FREE Full text] [doi: [10.2196/jmir.2966](https://doi.org/10.2196/jmir.2966)] [Medline: [24496109](https://pubmed.ncbi.nlm.nih.gov/24496109/)]
6. Griffiths KM, Calear AL, Banfield M, Tam A. Systematic review on Internet Support Groups (ISGs) and depression (2): What is known about depression ISGs? *J Med Internet Res* 2009;11(3):e41 [FREE Full text] [doi: [10.2196/jmir.1303](https://doi.org/10.2196/jmir.1303)] [Medline: [19793718](https://pubmed.ncbi.nlm.nih.gov/19793718/)]
7. McPherson M, Smith-Lovin L, Cook J. Birds of a Feather: Homophily in Social Networks. *Annu. Rev. Sociol* 2001 Aug;27(1):415-444. [doi: [10.1146/annurev.soc.27.1.415](https://doi.org/10.1146/annurev.soc.27.1.415)]
8. American Psychiatric Association. Diagnostic statistical manual of mental disorders (5th Ed.). Arlington, VA: American Psychiatric Publishing; 2013.
9. Newman MEJ. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences* 2006;103(23):8577-8582. [doi: [10.1073/pnas.0601602103](https://doi.org/10.1073/pnas.0601602103)]
10. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. 2009 Presented at: Third International AAAI Conference on Weblogs and Social Media; 17 May 2009; San Jose, CA.
11. Blondel V, Guillaume J, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J. Stat. Mech* 2008 Oct 09;2008(10):P10008. [doi: [10.1088/1742-5468/2008/10/P10008](https://doi.org/10.1088/1742-5468/2008/10/P10008)]
12. Lancichinetti A, Fortunato S. Community detection algorithms: a comparative analysis. *Physical Review E* 2009 Nov;80(5):6117. [doi: [10.1103/PhysRevE.80.056117](https://doi.org/10.1103/PhysRevE.80.056117)]
13. TechSmith. 2014. Jing Screen Recording Tool URL: <https://www.techsmith.com/jing.html> [accessed 2016-04-22] [WebCite Cache ID 6gwtSIXH9]
14. Iriberry A, Leroy G. A life-cycle perspective on online community success. *ACM Comput. Surv* 2009 Feb 01;41(2):1-29. [doi: [10.1145/1459352.1459356](https://doi.org/10.1145/1459352.1459356)]
15. Gruzd A, Haythornthwaite C. Enabling community through social media. *J Med Internet Res* 2013;15(10):e248 [FREE Full text] [doi: [10.2196/jmir.2796](https://doi.org/10.2196/jmir.2796)] [Medline: [24176835](https://pubmed.ncbi.nlm.nih.gov/24176835/)]
16. Young C. Community management that works: how to build and sustain a thriving online health community. *J Med Internet Res* 2013;15(6):e119 [FREE Full text] [doi: [10.2196/jmir.2501](https://doi.org/10.2196/jmir.2501)] [Medline: [23759312](https://pubmed.ncbi.nlm.nih.gov/23759312/)]
17. Newman MEJ. Mathematics of networks. In: Blume LE, Durlauf SN, editors. *The New Palgrave Dictionary of Economics*. 2nd Edition. Basingstoke: Palgrave Macmillan; 2008.
18. Chomutare T, Arsand E, Fernandez-Luque L, Lauritzen J, Hartvigsen G. Inferring community structure in healthcare forums. An empirical study. *Methods Inf Med* 2013;52(2):160-167. [doi: [10.3414/ME12-02-0003](https://doi.org/10.3414/ME12-02-0003)] [Medline: [23392282](https://pubmed.ncbi.nlm.nih.gov/23392282/)]
19. Nimrod G. The membership life cycle in online support groups. *International Journal of Communication* 2012;6(2012):1245-1261.
20. Carron-Arthur B, Ali K, Cunningham J, Griffiths K. From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities. *J Med Internet Res* 2015;17(12):e271 [FREE Full text] [doi: [10.2196/jmir.4705](https://doi.org/10.2196/jmir.4705)] [Medline: [26627369](https://pubmed.ncbi.nlm.nih.gov/26627369/)]

21. Nocera A, Ursino D. PHIS: A system for scouting potential hubs and for favoring their “growth” in a Social Internetworking Scenario. Knowledge-Based Systems 2012 Dec;36:288-299. [doi: [10.1016/j.knosys.2012.07.009](https://doi.org/10.1016/j.knosys.2012.07.009)]
22. Su W. Integrating and Mining Virtual Communities across Multiple Online Social Networks: Concepts, Approaches and Challenges. : IEEE Computer Society Press; 2014 Presented at: The Fourth International Conference on Digital Information and Communication Technology and its Applications; 6 May 2014; Bangkok, Thailand p. 199-204. [doi: [10.1109/DICTAP.2014.6821682](https://doi.org/10.1109/DICTAP.2014.6821682)]
23. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. J Med Internet Res 2012;14(3):e66 [FREE Full text] [doi: [10.2196/jmir.1854](https://doi.org/10.2196/jmir.1854)] [Medline: [22732103](https://pubmed.ncbi.nlm.nih.gov/22732103/)]

Edited by P Bamidis; submitted 20.07.15; peer-reviewed by V Waights, R Almeida, Q Zhang, S Yu, WC Su; comments to author 19.10.15; revised version received 06.12.15; accepted 07.03.16; published 30.05.16

Please cite as:

*Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM
Community Structure of a Mental Health Internet Support Group: Modularity in User Thread Participation
JMIR Mental Health 2016;3(2):e20
URL: <http://mental.jmir.org/2016/2/e20/>
doi: [10.2196/mental.4961](https://doi.org/10.2196/mental.4961)
PMID: [27242012](https://pubmed.ncbi.nlm.nih.gov/27242012/)*

©Bradley Carron-Arthur, Julia Reynolds, Kylie Bennett, Anthony Bennett, John Alastair Cunningham, Kathleen Margaret Griffiths. Originally published in JMIR Mental Health (<http://mental.jmir.org>), 30.05.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <http://mental.jmir.org/>, as well as this copyright and license information must be included.

Chapter 6

Foreword

This chapter presents the fourth of five empirical studies included in this thesis and the third of four that investigated an aspect of participation in BlueBoard followed by a comparative analysis of differentially-engaged users.

This study focused on the *content* created by users, applying an objective, automated machine-learning algorithm to identify topics based on the frequency of the co-occurrence of words in the board posts.

The study sought to answer the following questions:

RQ 5a: About what topics do BlueBoard users communicate?

RQ 5b: Are there systematic differences in the degree to which users with different levels of engagement write content pertaining to each of these topics?

RQ 5c: Do these differences change when higher-engaged users are communicating with lower-engaged users as opposed to communicating within their own grouping?

Unlike the community structure analysis in Chapter 5, the topics that were identified in this study reflected at least in part the disorder-specific structure of BlueBoard's sub-forums. However, it was found that the low-engaged users tended to write more content on these disorder-specific topics. The high-engaged users were found to be more likely than other users to write content resembling a consumer-defined perspective of recovery (see Chapter 1 discussion), but the focus was observed to shift to disorder-specific topics when responding to low-engaged users.

RESEARCH ARTICLE

Open Access



What's all the talk about? Topic modelling in a mental health Internet support group

Bradley Carron-Arthur^{*} , Julia Reynolds, Kylie Bennett, Anthony Bennett and Kathleen M. Griffiths

Abstract

Background: The majority of content in an Internet Support Group (ISG) is contributed by 1 % of the users ('super users'). Computational methods, such as topic modelling, can provide a large-scale quantitative objective description of this content. Such methods may provide a new perspective on the nature of engagement on ISGs including the role of super users and their possible effect on other users.

Methods: A topic model was computed for all posts ($N = 131,004$) in the ISG BlueBoard using Latent Dirichlet Allocation. A model containing 25 topics was selected on the basis of intelligibility as determined by diagnostic metrics and qualitative investigation. This model yielded 21 substantive topics for further analysis. Two chi-square tests were conducted separately for each topic to ascertain: (i) if the odds of super users' and other users' posting differed for each topic; and (ii) if for super users the odds of posting differed depending on whether the response was to a super user or to another user.

Results: The 21 substantive topics covered a range of issues related to mental health and peer-support. There were significantly higher odds that super users wrote content on 13 topics, with the greatest effects being for Parenting Role (OR [95%CI] = 7.97 [7.85–8.10]), Co-created Fiction (4.22 [4.17–4.27]), Mental Illness (3.13 [3.11–3.16]) and Positive Change (2.82 [2.79–2.84]). There were significantly lower odds for super users on 7 topics, with the greatest effects being for the topics Depression (OR = 0.27 [0.27–0.28]), Medication (0.36 [0.36–0.37]), Therapy (0.55 [0.54–0.55]) and Anxiety (0.55 [0.55–0.55]). However, super users were significantly more likely to write content on 5 out of these 7 topics when responding to other users than when responding to fellow super users.

Conclusions: The findings suggest that super users serve the role of emotionally supportive companions with a focus on topics broadly resembling the consumer/carer model of recovery. Other users engage in topics with a greater focus on experiential knowledge, disclosure and informational support, a pattern resembling the clinical symptom-focussed approach to recovery. However, super users modify their content in response to other users in a manner consistent with being 'active help providers'.

Keywords: Internet support group, Mental health, Topic modelling, Latent Dirichlet Allocation, Super users, Peer-support

Background

Online peer-to-peer communication is a popular source of health information and support. Recent research on Internet users in the USA found that 18 % of people had used the Internet in the last year to find information from a peer with similar health concerns [1]. Furthermore, 8 % of all Internet users had engaged in peer-support by either posting a question or sharing information based on their personal health experience [2]. Mental health concerns are a major component of this health information seeking,

with 28 % of all Internet users having sought mental health information online [3].

Given this popularity, there has been interest in determining whether Internet support groups (ISGs) are effective in reducing depressive symptoms. A systematic review of ISGs encompassing all types of health conditions failed to find convincing evidence that online peer-to-peer support was associated with a reduction in depressive symptoms [4]. Moreover, a review of depression ISGs specifically reported that there was a paucity of evidence concerning the effectiveness of depression Internet support groups for symptom reduction [5]. More recently, a randomised controlled trial of a depression ISG has provided

* Correspondence: bradley.carron-arthur@anu.edu.au
Centre for Mental Health Research, The Australian National University, 63 Eggleston Road, Acton, Canberra, ACT 2601, Australia

high quality evidence of depressive symptom reduction [6]. However, further research is required before firm conclusions can be drawn about the effectiveness of depression ISGs. It has been suggested that mental health internet support groups (MHISGs) increase the user's sense of empowerment [7] and such support groups are widely used in conjunction with other psychoeducational and therapeutic Internet interventions with the aim of promoting engagement [8]. More recently, a randomised controlled trial demonstrated that a depression ISG was associated with increased empowerment, self-esteem and perceived quality of life relative to a control condition [9].

Complementary to work focused on the effectiveness of MHISGs, another stream of research has been concerned with understanding the nature of ISGs [10]. Such understanding is vital to informing practice and policy to promote the growth and sustainability of ISGs [11]. Research on the nature of these peer-to-peer groups is also needed to identify what elements of the groups are responsible for fostering user empowerment, and what components might be enhanced to increase the potential effectiveness of ISGs for symptom reduction. In a series of studies on the nature of the Australian ISG BlueBoard, we have so far investigated the distribution of user engagement across the ISG [12], characteristics of users which predict user engagement and retention (Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A: User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study, Submitted) and the community structure of the ISG [13]. This research has shown that more highly engaged users: post vastly more than their peers in a distribution that follows Zipf's law (inversely proportional relationship between rank and frequency) [12]; tend to be consumers rather than carers (Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A: User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study, Submitted); tend not to be less than 20 years old (Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A: User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study, Submitted); and join earlier than the peers with whom they most often communicate, leading to the formation of sub-communities within the MHISG [13]. This research has highlighted the importance of peer-leaders who are highly engaged and who communicate with many other users. These findings are also consistent those from a previous survey of MHISG users which found that highly-engaged users identify themselves as 'active help providers' [14]. Based on a content analysis of user posts, it has been found that highly active users provide higher levels of social support than other users in the MHISG [15]. If social support underpins improvements in outcomes among users

of ISGs, the highly engaged user is likely to be an important contributor to the effectiveness of ISGs.

In a systematic review of studies investigating participation styles in online health communities, we found that the peer-leader phenomenon has been measured in a number of different ways in the literature [16]. However, the role of the peer leader has been most commonly operationalised as high posting frequency. For example, the top 1 % of users, labelled "superusers", have been observed to contribute around 75 % of all posts in the ISG [12, 17, 18]. In our review we noted that studies commonly attributed high value to the contributions of high-posting users despite the fact that a priori, posting frequency does not in of itself necessarily contribute value to the community [16].

In an attempt to develop a more nuanced index of post frequency that factored in post quality, Preece [19] recommended counting only posts which were "on-topic". This measure would appear to be preferable to an unadjusted frequency count. However, it assumes that posts can be validly dichotomised into 'on-' and 'off-' topic, a premise which is questionable in a mental health ISG where each person's lived experience and needs can vary, and the relevance of a post will depend on the perspective of the reader. It may be more helpful to conceptualise posts as being relevant to varying numbers of people and to measure how many people engage in each of the various topics. More particularly, given the large number of posts created by super users, there may be value in investigating if and how the topics and frequency with which they are discussed differ between super users and other users of the ISG, as well as comparing the responses of super users to fellow super users with their responses to other users. Identifying similarities and differences in the degree to which various topics are discussed by super users and others may indicate if the majority of post content is aligned with the interests of the majority of users. This may provide an indication of the role that super users are performing with respect to supporting other users. Thus, rather than asking, "What is a more accurate measure of peer-leadership than posting frequency", it may be more informative to ask, "Of what topics is the volume of posts in a mental health ISG comprised?" and, "How and in what circumstances does the frequency with which different topics are discussed vary between super users and other users?".

Content analysis

To date, most studies of the content of MHISGs have used human judges and pre-formulated coding schemes to manually classify different types of peer-support. A review of depression ISGs conducted in 2009 found that all eight of the studies which analysed content had used this method [10]. This includes typologies such as the Cutrona Support Behavior Code [20] and other systems [21, 22] which although differing somewhat generally involved common

categories such as: disclosure (revealing personal information about oneself); emotional support (showing empathy and concern, and offering affection and encouragement); companionship (engaging in activities with a person to communicate a sense of shared belonging); information support (providing helpful information); and cognitive guidance (advice, offering a new perspective from which to think about an issue).

In contrast to the manual methods outlined above, computer-aided methods have provided new ways of processing linguistic content which enable themes to be automatically and objectively detected in text on a large scale. One method which has been widely used is Latent Dirichlet Allocation (LDA) [23]. LDA is a topic modelling algorithm which determines latent topics across a corpus of text based on the distribution of words across the documents which make up the whole corpus. Words which co-occur in individual documents frequently across all the documents of the corpus are allocated to categories which represent a latent topic. LDA is an unsupervised machine learning method in which the algorithm derives the topics without using a training dataset. This contrasts with supervised machine learning in which an analyst teaches the algorithm to classify particular content. The two methods contribute different utility, with unsupervised methods such as LDA being particularly useful where the analyst is seeking to generate a summary of the data that is unbiased by human input, although the analyst may adjust some parameters such as the number of topics to be discerned. The output, referred to as *topics*, takes the form of groups of words ordered by their probability of occurring in that topic. The algorithm also computes the proportion of each document that is made up of each topic.

Several previous studies have used LDA to analyse MHISGs [24, 25] and peer-to-peer conversations about mental health in other online communities such as Twitter [26], Facebook [27] or other blogging sites [28]. These studies have shown that automated algorithms can be used to differentiate between the content of mental health specific communities and the content of other ISGs or online conversations [24, 26]. Furthermore, studies have successfully used the metrics obtained from these methods to make predictions about the diagnosis or symptom severity of users [27, 28]. One study [25] showed that it was possible to detect significant differences between the written content of highly-socially-connected users and other less-connected users. However, these studies were focussed on demonstrating the capability of the tools to differentiate between users. They did not analyse the nature of those differences in a manner that might increase our understanding of the social dynamics in a MHISG.

In the current study, we aimed to ascertain the predominant topics of discussion and investigate the nature of differences in content produced by super users

and other users of a MHISG. Our objective was to compare quantitatively the differences in frequency with which these groups of users write about various topics (as determined by LDA). We sought to compare the difference between the two groups as well as the difference between the responses made by super users in threads initiated by fellow super users compared with their response to threads initiated by other users.

Method

Data

Data for the current study was drawn from the log data of the Internet support group BlueBoard (blueboard.anu.edu.au). The dataset has been described previously [12, 13, (Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A: User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study, Submitted)]. Briefly, the data covered the period 1 October 2008 to 23 May 2014, in which 131,004 posts were made by 2932 users.

The Internet support group: BlueBoard

This service was provided by the Centre for Mental Health Research at The Australian National University. BlueBoard comprised 10 forums in which users communicated about a range of mental health issues, including: eight condition specific forums (in order of usage - depression, bipolar disorder, generalized anxiety, social anxiety, borderline personality disorder, eating disorder, panic disorder and obsessive compulsive disorder); one forum dedicated for carers of people with mental health issues; and one forum for general discussion. BlueBoard was first established in 2003 as a mood disorder group. It was closed in 2007 and 2008 due to lack of funding and re-established on 1 October 2008, but closed again on 30 June 2016. This second iteration of BlueBoard, from which the current data is drawn, did not include the content or registrations of previous users. Although BlueBoard provided a service, it was also designed with the aim of being used for research purposes. Moderators did not actively participate in any of the forums. Rather they enforced rules, for example, by editing posts to remove any personally identifying information, and alerting the infracting user via a private notification. Thus the data comprises content which is solely authored by BlueBoard users. BlueBoard otherwise includes features similar to other Internet support groups such as the ability to quote other users in posts and the provision user information (total posts and registration date) displayed beside the username of each post's author.

Analysis

Our analysis entailed two parts. In Part 1, a computer-aided content analysis was implemented using LDA in

order to model the predominant topics in BlueBoard post content. The output of this analysis included word lists which represented each of the identified topics as well as the proportion of words in each of the 131,004 posts that were from each topic. We converted these proportions to word frequencies in order to differentiate between posts with equal proportions but different word counts. In Part 2 of the analysis, we determined the degree to which the content of super users differed from that of other users in each of these topic categories.

Part 1: Identification of topics and sentiment

The LDA analysis [23] to model topics was implemented using Mallet software with hyperparameter optimization enabled [29]. The analysis was conducted using each post as a separate document over which the distribution of topics was to be determined. The standard Mallet stop-word list containing very common English words was used to exclude such words from the analysis. Additionally, 18 contractions of common words (e.g. you're) and the words 'thing' and 'things' were excluded.

Our objective in developing a topic model was to provide an intelligible representation of the type of content BlueBoard comprises. We therefore sought to model topics in the corpus at a level of granularity that was neither too broad nor too specific. Accordingly, selection of an appropriate number of topics was carried out by author BC following a method developed for this purpose by Evans [30]. We implemented the LDA analysis iteratively, modelling between 10 and 100 topics in increments of 10.

In the first of three phases of the analysis, we investigated the intelligibility of these models based on qualitative investigation of the word lists comprising each topic. In models containing more than 30 topics, we observed some duplications of topics where the meaning of the keywords was overly similar. For example, in the model containing 40 topics, the topic Livelihood split into two topics with top ranked words "Money, time, pay, day" in one and "work, job, week, working" in the other. This degree of redundancy in a model with more topics was judged to be less intelligible. Below 20 topics, we observed mergers of clearly distinguishable topics which were distinct in models with higher numbers of topics. For example, the topics Therapy and Livelihood were observed together in a single topic with top ranked words including "work, job, psych, health".

In the second phase of the LDA analysis, we incorporated the quantitative measures of each topic's coherence [31] and specificity (distance to corpus score) [32] provided by the Mallet software diagnostic output. We used these in models of 20, 25 and 30 topics to flag topics which initially seemed substantive, but which were possibly incoherent or similar to a representation of the entire corpus.

In the third and final phase of the LDA analysis, we used domain knowledge of the context in which the topics

occurred most frequently and the posts in which the topics were highly prevalent to validate their nature. Based on this combination of qualitative, quantitative and domain knowledge, we concluded that a model with 25 topics was the most intelligible. However, there were four 'junk' topics among these for which a meaningful interpretation could not be identified. This included two topics which had low coherence scores relative to the other topics (eg, "back head pain put water body eyes cold cat front") and two topics which had low specificity (distance to corpus scores) (eg, "back told time thought home felt day friend wanted asked"), although coherence and specificity scores were not used as executive criteria for exclusion. After excluding the four meaningless topics, the final model included 21 topics. Labels were assigned to each topic by BC with the assistance of the MHISG manager (JR) who provided domain expertise based on the overarching concept apparent in each topic's word list and by perusing posts that were comprised predominantly of each single topic.

Part 2: Comparison of topic and sentiment expression across user groups

Users were divided into "super users" (the top 1 % of users by posting frequency) and "other users" (the remaining 99 % of users). As shown in Table 1, the 29 super users contributed just under three times as many posts in total as the 2903 other users, but their individual posts contained significantly fewer words (a difference of 35 words between medians; Mann-Whitney $U = 1.23 \times 10^9$, $p < .001$). In total, other users initiated more than twice as many threads as super users. Super users contributed more than four times as many posts to the threads that were initiated by fellow super users than to the threads of other users. Other users contributed three times as many posts to threads initiated by fellow other users than to threads of super users.

To determine the nature of the content contributed by super users we conducted two chi-square tests for each of the 21 topics. In the first, we analysed all posts, comparing the odds that the two different user group's posts contained content from each topic. In the second, we analysed only super user posts. We similarly compared the topic-specific odds of posts written by super users in response to super users with those that super users wrote in response to other users. We used a Bonferroni correction

Table 1 User groups

	Super users (n = 29)	Other users (n = 2903)
Total number of posts	96,896	34,108
Total threads initiated	2133	4607
Posts in super user threads	79,584	8,358
Posts in other user threads	17,312	25,750
Mean (sd) post word count	70 (114)	110 (143)
Median post word count	35	67

to adjust for multiple comparisons ($n = 42$). Thus a value of $p < .001$ was required for statistical significance.

Results

Part 1: Identification of topics

The 10 most frequently occurring words in each of the 21 topics produced by the LDA analysis are shown in Table 2. The topic comprising the largest percentage of the corpus was Social Relations, making up 8.21 % of all words in the collection (excluding stop words).

The distinctiveness of several topics was apparent directly from the word lists output by the LDA analysis. Their nature was evident from the presence of multiple words with a common theme. These topics were Medication, Therapy, Livelihood, Entertainment, Family, BlueBoard, Food and Drink, Affection, Bipolar, Anxiety, and Drugs and Alcohol. Table 3 presents a short quotation from the corpus to illustrate each of these topics.

The nature of other topics was more apparent once the words were contextualised by perusing posts for which the topic made up the majority of the content and by incorporating the domain expertise of the BlueBoard manager. Table 3 presents quotations that provide the context for interpreting these topics.

Three topics (Parenting Role, Co-created Fiction and Chat) almost exclusively involved super users. The usernames that were included in the word lists for these topics were those of the super users who were known to regularly engage in conversations on these topics. Parenting Role largely comprised conversation about managing one's parental responsibilities while also managing one's mental illness. Co-created Fiction comprised words and usernames which were frequently included in stories narrated in the third person by a particular sub-group of super users. The stories were imaginary journeys in which the users underwent heroic challenges and supported each other. Chat comprised words that were highly typical of conversation in the largest thread on BlueBoard – “The Beer Garden”. This was a user-established thread which users visited for companionship and to socialise.

The topic Mental Illness comprised meta-level content referring to mental illness in general terms. It differed from the topics Bipolar, Anxiety and Depression which comprised content which was more specifically about the individual's personal experience of these illnesses. The topic Depression was particularly exemplary of this, comprising many words that were not semantically related, but which were experientially related to depression. The topic Depression had the lowest specificity (distance to corpus score) of all the topics, signifying that of all the topics it was the most similar to a random selection of words from the whole corpus. The most frequently occurring word was depression itself. However, based on the other words in the list, it appears that the topic comprised a broader

Table 2 The 10 most frequently occurring words in each topic

Topic Label	Proportion of corpus (%)	Top 10 most frequent words
Social Relations	8.21	people person make feel good time understand life relationship wrong
Depression	7.20	depression feel people friends time years life anxiety work year
Anxiety	7.09	feel feeling anxiety bad time hard head thoughts sick day
Recovery Journey	6.11	life love feel pain world heart time hope find day
Daily Functioning	5.09	day sleep night today good work bed time morning tomorrow
Affection	5.01	hugs hope big good love thinking xxx hear hug happy
Family	4.83	family daughter kids mum mother son husband parents time children
Therapy	4.56	health mental good support psychologist therapy find talk psych psychiatrist
Medication	4.32	meds medication anxiety taking side depression effects weeks doctor dose
Mental Illness	4.13	people mental illness bipolar good depression disorder life problems important
BlueBoard	4.03	post read thread write writing posts squad mod people reading
Positive Change	3.34	life time years happy great love wonderful world part important
Co-created Fiction	2.85	[username] ^a eyes back [username] ^a head hand face tears water ship
Livelihood	2.79	money work job people pay Centrelink ^b government system Australia health
Chat	2.72	lol beer dog cool awesome love gonna yeah [username] ^a [username] ^a
Parenting role	2.63	kids bit [username] ^a time school life stuff kinda mum daughter
Bipolar	2.50	bipolar depression mood disorder manic diagnosed meds normal mania diagnosis
Philosophy	1.97	world brain mind people god human women science society power
Entertainment	1.61	music love song play movie good watch book playing songs
Food and Drink	1.58	eat food tea eating water chocolate coffee good chicken drink
Drugs and Alcohol	1.14	alcohol drink drinking smoking shit drugs smoke drug stop weed

^aIn accordance with the study ethics protocol usernames have been omitted

^bCentrelink is an Australian Government organisation which provides social services and welfare payments

notion of depression than the clinical definition, encompassing themes of dysphoria and social isolation which may be common among people seeking social support.

Table 3 Quotations exemplifying the nature of each topic

Topic	Exemplary quotations
Social Relations	"I can see now that I have been in controlling relationships but I'm not weak enough to just accept it & not strong enough to stand up for myself enough." "In life friends come and go, its the natural order, and if people are in a better place then it is a good thing."
Depression	"Um, hi. I'm new to this website, so I just wanted to share a bit about myself. I've dealt with depression for almost four years of my life now (I am 18) and also have social anxiety. I have no one to talk to in my life because most of my friends are all too busy for me apparently... Anyway, I could go on forever about my depression and my story, but I won't bore you guys. (just yet)"
Anxiety	"I've always struggled with stress & anxiety, and have had mild panic attacks in the past. But what I have tended to do lately is be 'triggered' by something, then get really emotional, cry, stress out, non stop thoughts, at time weird breathing etc."
Recovery Journey	"I am strong because I am weak. I am beautiful because I know my flaws. I am a lover because I'm a fighter. I am fearless because I have been afraid. I am wise because I have been foolish... And I can laugh because I've known sadness" "Hi [username] My thoughts and feelings are very much the same, but i was put on this earth for a reason and I'm sure it was not to suffer for my whole life, there is happiness out there I believe that otherwise what is the point of being here"
Daily Functioning	"Hi guys, Had a pretty lame day today. Because I got so little sleep last night (~3 h), I ended up getting my daughter an extra day at childcare. Then I was supposed to be productive, but got almost nothing done all day. My anxiety has really "re-generalised" lately... Will see how I go with work tomorrow... I'm not feeling too bad really despite today... Just feeling tired... More sleep tonight will help"
Affection	"Hi sweetie, Big hugs to you, sounds like a hard day Know what, I don't think we've been introduced...I'm [username], welcome. Can't imagine how tough it would be, hang in there, sending hugs and happy vibes your way."
Family	"It pains me to read about families that do not support each other! I realize i have been fortunate to be raised by loving parents and be able to become a good parent to my now grown children. I love to see my children support each other and continue the close bonds that have formed over the years. I wish this was possible for all families."
Therapy	"I really find it hard to talk in therapy. I just feel that the therapist does not feel the same way as I do so there is no way he can relate. That is why I decided to come here."
Medication	"Hi I have anxiety (OCDs) and depression. I am currently on several medications - 1 week ago I increased Zoloft (Sertraline) to 200 mg, prior to that 50 mg - 150 mg over 4 weeks. My question is has anyone on BB had experience with Zoloft and know how long it takes to kick in?"
Mental Illness	"Reading as you do, [username], will help you on your journey. I have been doing what you are doing now for many years. I have changed my mind many times and considered the biological, environmental and psychological forces of what we call bipolar. "Dear [username], you are right. People with a mental illness are under represented in research. Not many of us are able to do research on mental illness. There are many reasons for this"
BlueBoard	"Hey [username] :) Great to see you posting again. To find the thread about new avatars, go to the Blueboard Notices forum, Sub-forum Blueboard Notices, thread titled Image and Photo Posting. The last post on that thread explains how to upload custom avatars."
Positive Change	"I am deeply proud to be a member of this wonderfully diverse community of individuals, linked by mental health issues and yet so different in life circumstances. We all have so much to offer each other by sharing our lives, our trials and triumphs. I am a better person for having the good fortune of being introduced to you all through BlueBoard. Humbly and wholly, I offer up my thanks to you all" "Truth be told, my friend, I would not want you to take away my pain. It is an important motivator in my life, an important process. I do not identify myself through this pain, but it acts as a catalyst for change in my life. From this great perceived negative, I am reborn into the positive"
Co-created Fiction	"[username] gives out an audible cry of fear, [username] is instantly at her side, arm around her, dagger in hand. She hugs [username], eyes fixed on the flaming bird of destruction which is almost upon them. The crew shouts increase as they prepare the cannons and water down the ship" "Tears fall onto her lap like a waterfall, she hides her face in her hands sinking back to the wet wood soaked from her tears."
Livelihood	"Hi [username], if you have some financial trouble and need some money to fix your car there are government organization who can lend you some interest free money."
Chat	"We should have beer garden day. lol everyone sitting around in their backyard beer gardens or out at beer gardens" "[username] is BACK!!!!!! Wooooooo hoooooo!!!!!! :D :D Hasn't been the same without you! This calls for an undies-on-head dance! [username]!!! Get thee to the Beer Garden STAT!!!
Parenting role	"School excursion tomorrow. daughter is soooooo excited. oh crap- i didnt organise everything- usually i do that on Thursday- but i should have done that tonight- after school snack, footy stuff and piano stuff as well as breakkie and getting kids ready." "Im sure those routines will need adjusting etc.- but its kinda what we do, try to do or what we want to do. so... thinking a bit of time management might make it all happen more. and it fits around after school things"
Bipolar	"For me it varies quite a lot. I can go lengthy periods where everything is fine and episodes are infrequent though long when they do hit. At other times, I cycle more rapidly and can go from depression to mania with no balanced state in between."
Philosophy	"Physics postulates that there are an infinite number of parallel universes in existence, all of them either subtly or vastly different from each other. Psychology and Philosophy confirm this, every one of us is a parallel universe of subjective reality, uniquely coexisting with the others." "In it's simplest form, this is kind of my thesis: In societies dominated by patriarchal attitudes, cautionary tales often depict powerful women as physically and emotionally ugly... which greatly disadvantages women in society striving for power"

Table 3 Quotations exemplifying the nature of each topic (Continued)

Entertainment	"The song in free to listen to on uT on the net. Most of his songs are. Tony Joe White is one of my favorite singers."
Food and Drink	"I was wondering if anyone here enjoys cooking? I am trying to make meals which have more vegetables and healthy foods"
Drugs and Alcohol	"im so blerrrr, spent so much, trying to keep the boredom away too much booze too much smoke just to get myself out of bed i had to dilute myself, deression is hell"

Part 2: Differences between user groups in frequency of topic expression

(i) Super user vs other users: across all posts

The total number of words written by each user group on each topic is shown in Fig. 1. It can be seen that for all topics other than Depression and Medication, super users wrote more content than other users.

To investigate the relative inclinations of users to write content on each of these topics, we compared the odds that their content was 'on' versus 'not on' each topic. The results are shown in Table 4. Super users wrote relatively less than other users on seven topics: Depression, Medication, Therapy, Anxiety, Bipolar, Daily Functioning and Social Relations. For example, the odds that super users incorporated content containing words from the topic Depression were one-quarter that of the odds that other users incorporated such content. Conversely, the odds that super users wrote content were significantly higher for 13 topics. These were Parenting Role, Co-created

Fiction, Mental Illness, Positive Change, Chat, Philosophy, Affection, Entertainment, Recovery Journey, Food and Drink, BlueBoard, Livelihood, and Family. The largest discrepancy was for the topic Parenting Role, with the odds of super users posting on this topic being 7.97 times that for other users. There was not a significant difference in posting by the two user groups for the topic Drugs and Alcohol.

(ii) Super user responses to super user vs other user threads (posts by super users only)

The number of words written by super users on each topic in response to fellow super users compared with responses to other users is shown in Fig. 2. On each topic, super users wrote more content in response to fellow super users than other users. Results of chi square tests comparing the odds of response for the two types of responses are shown in Table 5 for each topic. Odds were significantly lower for responses to super users compared to responses to other users for seven topics: Therapy, Medication, Bipolar, Depression, Social Relations, Mental Illness,

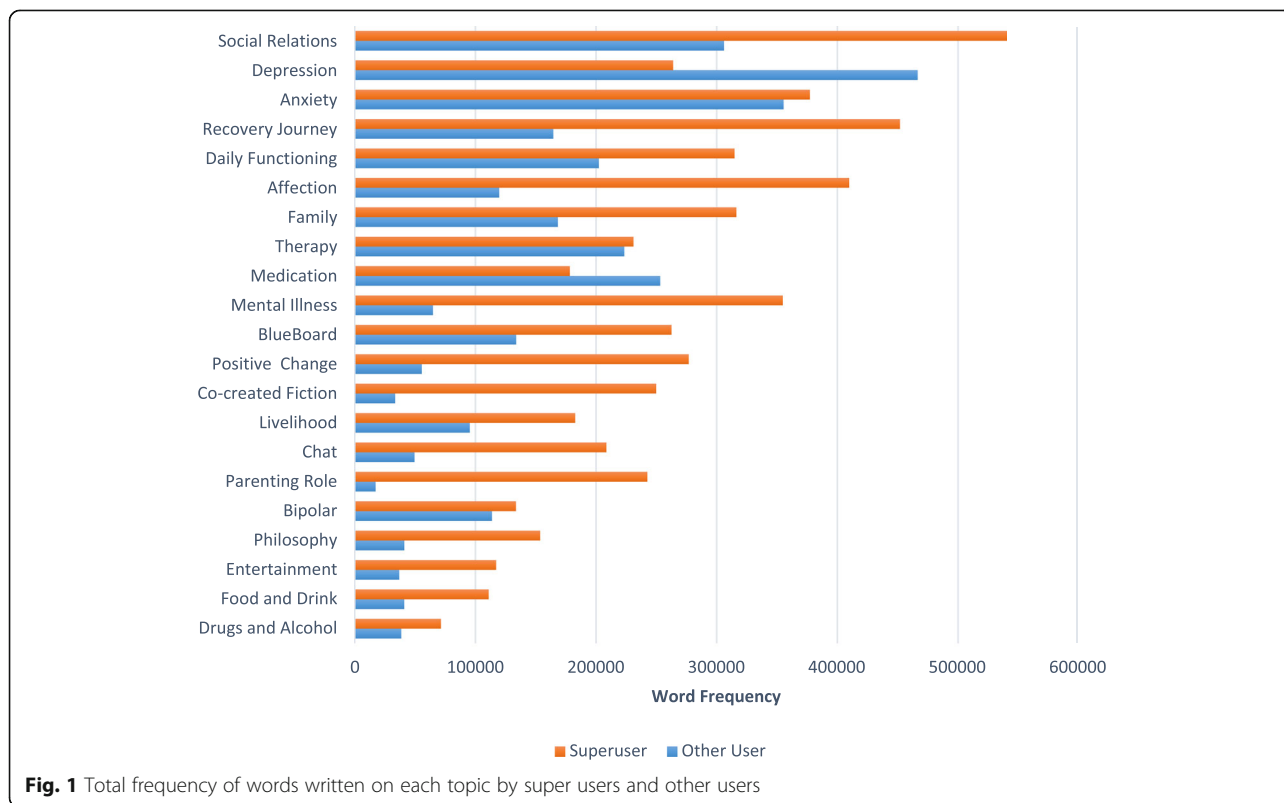


Fig. 1 Total frequency of words written on each topic by super users and other users

Table 4 Odds ratios and chi-square analyses for super user vs other user content for each topic

Topic	Chi Square	Odds Ratio (95 % CI)	P
Depression	284428.58	0.27 (0.27–0.28)	<0.0001
Medication	108165.68	0.36 (0.36–0.37)	<0.0001
Therapy	40040.97	0.55 (0.54–0.55)	<0.0001
Anxiety	60723.13	0.55 (0.55–0.55)	<0.0001
Bipolar	12566.17	0.63 (0.63–0.64)	<0.0001
Daily Functioning	3423.64	0.84 (0.84–0.85)	<0.0001
Social Relations	239.74	0.96 (0.96–0.97)	<0.0001
Drugs And Alcohol	4.52	1.01 (1.00–1.03)	0.03
Family	87.00	1.03 (1.02–1.04)	<0.0001
Livelihood	147.06	1.05 (1.04–1.06)	<0.0001
Blueboard	460.79	1.08 (1.07–1.08)	<0.0001
Food And Drink	4755.95	1.49 (1.47–1.51)	<0.0001
Recovery Journey	21845.44	1.55 (1.54–1.56)	<0.0001
Entertainment	8999.79	1.76 (1.74–1.78)	<0.0001
Affection	40184.47	1.94 (1.93–1.96)	<0.0001
Philosophy	17719.14	2.08 (2.05–2.10)	<0.0001
Chat	30423.66	2.35 (2.33–2.38)	<0.0001
Positive Change	52598.04	2.82 (2.79–2.84)	<0.0001
Mental Illness	76564.00	3.13 (3.11–3.16)	<0.0001
Co-Created Fiction	70981.96	4.22 (4.17–4.27)	<0.0001
Parenting Role	96582.60	7.97 (7.85–8.10)	<0.0001

and Positive Change. The largest discrepancy was for Therapy, which showed that super users were one-third as likely to write about therapy in response to super users compared to other users. The odds that responses to super users contained more content were significantly higher for 13 topics. These were Co-created Fiction, Philosophy, Chat, Entertainment, Food and Drink, Livelihood, Recovery Journey, Daily Functioning, Drugs and Alcohol, Parenting Role, Anxiety, BlueBoard, and Affection. The largest discrepancy was for the topic Parenting Role, with super users being 5.29 times as likely to write about this content in response to super users than in response to other users. There was not a significant difference for the topic Family.

Discussion

Principal findings

The current study used a computer-assisted method to identify topics in a mental health Internet support group. In particular, an analysis using LDA enabled us objectively to identify 21 topics which constituted the major components of discussion on the mental health ISG BlueBoard. There were significant differences in the frequency with which highly engaged super users wrote content on these topics compared to other users and between the content of super user posts in response to posts from users of each group. The pattern of results was consistent with a model that suggests more highly engaged users play a role as active help providers,

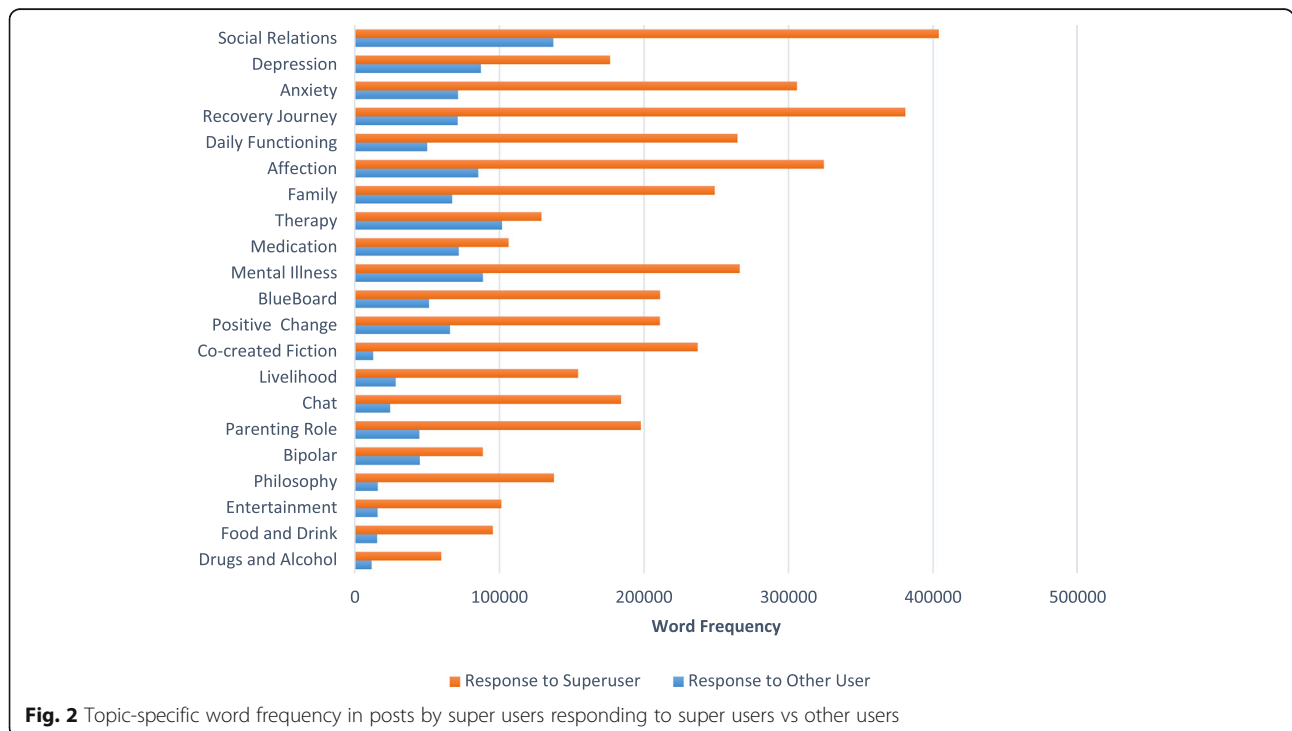


Fig. 2 Topic-specific word frequency in posts by super users responding to super users vs other users

Table 5 Odds ratios and chi-square analyses comparing super user responses to super users vs other users across topics

Topic	Chi Square	Odds Ratio (95 % CI)	p
Therapy	74299.69	0.32 (0.32–0.33)	< .0001
Medication	39546.97	0.39 (0.38–0.39)	< .0001
Bipolar	12380.59	0.53 (0.52–0.53)	< .0001
Depression	22647.52	0.53 (0.53–0.54)	< .0001
Social Relations	5770.87	0.78 (0.77–0.78)	< .0001
Mental Illness	2961.97	0.80 (0.80–0.81)	< .0001
Positive Change	1026.92	0.86 (0.86–0.87)	< .0001
Family	0.75	1.00 (1.00–1.01)	0.39
Affection	70.18	1.03 (1.03–1.04)	< .0001
Blueboard	522.38	1.12 (1.11–1.13)	< .0001
Anxiety	1417.75	1.18 (1.17–1.19)	< .0001
Parenting Role	1304.21	1.21 (1.20–1.23)	< .0001
Drugs And Alcohol	1105.30	1.40 (1.37–1.43)	< .0001
Daily Functioning	5892.94	1.46 (1.45–1.48)	< .0001
Recovery Journey	9230.50	1.50 (1.49–1.51)	< .0001
Livelihood	3861.81	1.50 (1.48–1.52)	< .0001
Food And Drink	3687.71	1.69 (1.66–1.72)	< .0001
Entertainment	4388.67	1.76 (1.73–1.79)	< .0001
Chat	12012.34	2.09 (2.07–2.12)	< .0001
Philosophy	11360.35	2.39 (2.35–2.43)	< .0001
Co-Created Fiction	41183.62	5.29 (5.19–5.38)	< .0001

particularly with respect to the provision of companionship and emotional support, but are relatively less inclined than other users to write content about clinical topics such as medication and treatment.

Part 1: Identification of topics

The topics identified in the current study comprised meaningful themes with relevance to mental health. The most frequently discussed topic was Social Relations, followed by two condition-specific topics, Depression and Anxiety. Bipolar was also identified as a topic. These three condition-specific topics are also the subject of disorder specific sub-forums on BlueBoard, e.g. “Living with Depression”, which contain the highest number of posts of all disorder specific sub-forums. This lends some support to the validity of the LDA analysis. It is important to note that these topics and sub-forums are not redundant categorisations. This is because BlueBoard sub-forums often contain content that does not pertain specifically to a single disorder. For example, content about experience of comorbid anxiety and depression may be found in either or any forum. The LDA topics are able to distinguish between these types of content, regardless of their location in the organisational structure of BlueBoard.

From the perspective of the framework proposed by Preece [19], some of the identified topics may be classified as ‘off-topic’, for example Chat, whereas other topics may be classified as ‘on-topic’, for example Therapy. This framework is limited in that it fails to consider the relevance of ‘off-topic’ content within the context of its broader utility. The findings of the current study are more consistent with frameworks which classify different types of content into categories of peer-support. From this perspective, the topic Chat may be considered relevant to companionship, and the topic Therapy may be considered relevant to informational support. Thus, although Chat is not directly related to mental health, its relevance in the context of peer-support is apparent.

There are clear links between the topics identified by automated processing in the current study and the different types of peer-support content that have previously been defined in the social-support literature [33] and implemented in content analyses of MHISGs [15]. In particular, the topics identified in the current study were consistent with the social-support categories of disclosure, experiential knowledge, information support, companionship, emotional support, group structure and process, and cognitive guidance. Topics involving specific types of mental health issues including Depression, Anxiety, Bipolar, and Drugs and Alcohol, comprised users’ first and second hand experiences. Thus, they could be seen to be related to ‘disclosure’ and ‘experiential knowledge’; e.g. “hi. I’m new to this website, so I just wanted to share a bit about myself. I’ve dealt with depression for almost four years of my life now”. Topics involving circumstances and contexts in which mental health issues occur and impact, including Daily Functioning, Social Relations, Livelihood, Family, Food and Drink, and Parenting Role, also fit disclosure and experiential knowledge categorisations; e.g. “I become so depressed I cant get out of bed, dont eat, cry all the time..... I just cant function” . The topics Medication and Therapy may also fit these categorisations but given their direct role in treatment, it is likely that their primary relevance is to the peer-support category of ‘information support’; e.g. “I’ve been using lamictal and epilium for a number of years, it’s the best combination I’ve come across.”. The topics Chat, Entertainment and Philosophy involved conversational matters typical of ‘companionship’; e.g. “I may be feeling low...but this gave me a laugh, it’s my type of humor exactly...” . The topic Affection related to care and concern for others, factors that are typical of ‘emotional support’; e.g. “Can’t imagine how tough it would be, hang in there, sending hugs and happy vibes your way.”. The topic Co-created Fiction included elements of both companionship and emotional support; e.g. “[username] gives out an audible cry of fear, [username] is instantly at her side, arm around her, dagger in hand.”. Finally, the topic BlueBoard involved references to the forum itself and thus concerned

references to ‘group structure and process’; e.g. “Not sure if you read post #116 in this thread - we posted stuff at the same time, so you may have missed it.”

The remaining topics, Positive Change, Recovery Journey and Mental Illness are consistent with ‘cognitive guidance’ though not necessarily in the form of direct advice. All three topics included the word “life” and it is clear from our perusal of typical posts that the content explores the impact of mental illness in their life and the experience of finding a pathway to recovery; e.g. “I know many people with bipolar who are happy with their life no matter what problems they face. Yet what our world teaches people with mental illness is that they are unlucky and that their life will be a troubled one”. This content is consistent with the consumer/carer model of recovery which emphasises personally meaningful elements of recovery such as hope, healing, empowerment and connection [34]. This contrasts with the traditional clinical model which focuses on the efficacy of treatments in reducing the symptoms that formally define mental illnesses. This clinical symptom-focussed approach is more apparent in topics which focus directly on illnesses and treatments e.g. Bipolar and Medication.

We highlight the above associations between the LDA topics and social support content for the purpose of describing trends observed in the data across user groups, which are discussed below. We acknowledge these above associations and below trends represent our interpretation of the data and that this interpretation is subjective. However, our methodology has largely deferred the point at which subjective interpretation enters the study until after the computation of the results (with the exception of the number of topics selected). Thus the strength of this analysis is not only in the novel perspective it provides on quantitative large-scale trends in the data, but it is also inherent in the transparency and replicability of the analysis. A content analysis of this nature has not previously been conducted on a MHISG.

Part 2: Differences between user groups

Overall, compared to other board users, super users were relatively more engaged in topics which related to companionship, emotional support and cognitive restructuring with a focus on consumer/carer defined recovery. This was the case for 10 out of the 13 topics for which super users wrote relatively more content than other users. In contrast, other users were relatively more engaged in topics which related to disclosure, experiential knowledge and informational support. This was the case for all seven of the topics for which other users wrote relatively more content than super users. This suggests a greater focus by other users on the traditional clinical symptom-focused approach to recovery.

Although super users have been known to identify themselves as ‘active help providers’ [14], the current study raises the question of whether the type of support provided by these users matches the type of support being sought. The findings of the current study suggest that there is an overall discrepancy in the type of content in which the two user groups prefer to engage. If this difference is the consequence of a discrepancy in perceptions of recovery, super users may be well placed to support fellow super users, but not necessarily be the best placed to provide the support being sought by other users. However, it is also evident that super users change the nature of their content depending on type of user to whom they are responding. Super user responses in other users’ threads were found to have higher odds for 5 of the 7 topics in which other users were previously observed to be relatively more engaged, and in particular in topics that were typical of experiential knowledge and informational support such as depression and medication. This suggests that super users actively change the type of content they contribute to align with the interests of other users when they are responding to them. This is consistent with the idea that super users are generally ‘active help providers’. However, we acknowledge that there may be other explanations for the data and that responsiveness to communication context does not necessarily imply that the super user is delivering help. For example, based on the current data, we cannot exclude the possibility that they are seeking or receiving help. Further research is required to investigate this issue.

A previous qualitative investigation of BlueBoard that examined depression information needs [35], found that ‘coping with depression’ (in particular, symptoms) was the most frequently explicitly and implicitly stated information need in user posts. This finding appears consistent with the type of content most often posted by other users in the current study. Furthermore, in another study of BlueBoard which involved a qualitative analysis of user-perceived advantages of participation in the board, the two most frequently cited types of advantage were (i) positive personal change, encompassing emotions such as: feeling glad, grateful, hopeful and inspired; cognitive effects, including changing the way a person thought about an issue; and behavioural effects, including choosing to see a doctor; and (ii) positive social interactions and support [36]. Few posts referred to symptom or disorder-specific advantages and it was concluded that the benefits of forum participation may be best conceptualised in terms of supporting overall recovery rather than as disorder or symptom-specific effects. In concert with the findings of the current study, this may suggest that the input provided by super users is consistent with the type of support that is valued by members, and is also consistent with broader evidence regarding the benefits

of participating in ISGs, including increased sense of empowerment [7, 9], and improved perceived quality of life [9] and self-esteem [9].

Limitations and future research

The current research has three main limitations. The frequency with which topics are expressed in a MHISG was represented in the current study in both absolute and relative terms. However, such data does not provide insight into the subjective experience of the user reading the topic content. For example, a post may contain a short statement about medication followed by a long story about a person's experience with depression. The reference to the medication may be of great importance in the story, but due to the associated small word count the subjectively important status is not represented in the data. Thus the current research methodology can provide a broad overview of the nature of the content in a MHISG, but is limited in the extent to which it represents a user's intention or another user's interpretation of the content. Future research seeking to address this may incorporate qualitative and or supervised machine learning methods to the analysis [37].

The second limitation of the study is that the dichotomisation of users according to their posting frequency (super user vs other users) provides a limited framework for defining the nature of a user's participation in an ISG. The role performed by a user can be classified by more nuanced metrics [16] and may change over time [38]. However, posting frequency is the most common way that research has classified users to date [16], often with an assumption that users with higher frequencies of posts contribute greater value to the ISG. The current study demonstrates broadly that these users contribute *different* value. Future research may demonstrate further differentiation in the kinds of value contributed by different users. In non-mental health contexts, researchers have focused on differentiating providers and seekers of emotional support, information support and companionship in ISGs [38–40]. In a MHISG context there may be other important roles such as users whose posts facilitate a decrease in self-stigma or users who are effective in supporting other users who are in crisis to seek professional care.

Lastly, it is both a strength and limitation of this study that it was focused on a MHISG in which moderators do not play an active role in the community. Consequently the results may not generalise directly to differently constituted MHISGs. However the findings are strengthened by the fact that the behaviour we have observed occurred without the potentially biasing influence of ISG staff. Further research is required to understand how this may vary in different MHISGs

Conclusion

The current study demonstrates the utility of a computational method for analysing the content of MHISGs. This technique enables trends in user engagement patterns to be investigated objectively and on a large scale. The pattern of findings in the current study has provided support for the notion that the most active members in a MHISG are generally 'active help providers'. The findings suggest that super users serve the role of emotionally supportive companions with a focus on topics broadly resembling the consumer/carer model of recovery. Other users engage in topics with a greater focus on experiential knowledge, disclosure, and informational support, a pattern resembling the clinical symptom-focussed approach to recovery. However, super users also modify their content to be more like that of other users when responding to them. These findings highlight similarities between the nature of super user engagement and existing evidence regarding the therapeutic outcome of user participation in ISGs, suggesting that the most highly engaged users may play an important role in this outcome.

Abbreviations

ISG: Internet support group; LDA: Latent Dirichlet Allocation; MHISG: Mental health internet support group

Acknowledgements

We are grateful for the general support provided by John Alastair Cunningham during the progress of this study.

Funding

B. Carron-Arthur is supported by an Australian Postgraduate Award. K.M. Griffiths is supported by the Australian National Health and Medical Research Council (NHMRC) Research Fellowship 1059620. BlueBoard is supported by funding from the Australian Department of Health.

Availability of data and materials

The data supporting the conclusions of this article is publicly available via blueboard.anu.edu.au.

Authors' contributions

BC conceived the study, undertook the analyses and wrote the paper. KB and KG contributed to the study design. JR and KG contributed to the interpretation of the data. JR, AB and KB contributed to the data collection. All authors edited the paper and have read and approved the final manuscript.

Competing interests

K.M. Griffiths established BlueBoard; and K.M. Griffiths, J. Reynolds, K. Bennett, and A. Bennett are responsible for the provision of the BlueBoard service. None of the authors derives personal financial benefit from the operation of the service.

Consent for publication

The consent to publish individual's data, with usernames removed, was informed and obtained through an online click wrap agreement.

Ethics approval and consent to participate

The research has been approved by the Australian National University Human Research Ethics Committee. The consent to participate was informed and obtained through an online click wrap agreement.

Received: 18 May 2016 Accepted: 17 October 2016

Published online: 28 October 2016

References

1. Fox S. Peer-to-peer Health Care. Washington: PewResearchCenter; 2011.
2. Fox S, Duggan M. Health Online 2013. Washington: PewResearchCenter; 2013.
3. Fox S, Jones S. The social life of health information. Washington: PewResearchCenter; 2009.
4. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Br Med J*. 2004;328(7449):1166. PMID: 15142921.
5. Griffiths KM, Calear AL, Banfield M. Systematic Review on Internet Support Groups (ISGs) and Depression (1): Do ISGs Reduce Depressive Symptoms? *J Med Internet Res*. 2009;11(3):e40. PMID: 19793719.
6. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The Effectiveness of an Online Support Group for Members of the Community with Depression: A Randomised Controlled Trial. *PLoS One*. 2012;7(12):e53244.
7. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. *Comput Hum Behav*. 2008;24(5):1867–83.
8. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries N, Brug J, et al. Which Intervention Characteristics are Related to More Exposure to Internet-Delivered Healthy Lifestyle Promotion Interventions? A Systematic Review. *J Med Internet Res*. 2011;13(1):e2.
9. Crisp DA, Griffiths KM, Mackinnon AJ, Bennett K, Christensen H. An online intervention for reducing depressive symptoms: Secondary benefits for self-esteem, empowerment and quality of life. *Psychiatry Res*. 2014;216(1):60–6.
10. Griffiths KM, Calear AL, Banfield M, Tam A. Systematic Review on Internet Support Groups (ISGs) and Depression (2): What Is Known About Depression ISGs? *J Med Internet Res*. 2009;11(3):e41.
11. Young C. Community management that works: how to build and sustain a thriving online health community. *J Med Internet Res*. 2013;15(6):e119. PMID: 23759312.
12. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions*. 2014;1(4):165–8. doi:10.1016/j.invent.2014.09.003.
13. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM. Community structure of a mental health Internet support group: modularity of user thread participation. *JMIR Ment Health*. 2016;3(2):e20.
14. Nimrod G. The membership life cycle in online support groups. *Int J Commun*. 2012;6:1245–61.
15. Salem D, Bogat G, Reid C. Mutual help goes on-line. *J Community Psychol*. 1997;25(2):189–207.
16. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities. *J Med Internet Res*. 2015;17(12):e271.
17. van Mierlo T. The 1% rule in four digital health social networks: an observational study. *J Med Internet Res*. 2014;16(2):e33. PMID: 24496109.
18. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. *J Med Internet Res*. 2012;14(3):e66. PMID: 22732103.
19. Preece J. Sociability and usability in online communities: determining and measuring success. *Behav Inf Technol*. 2001;20(5):347–56.
20. Cutrona SC, Suhr J, MacFarlane R. Interpersonal transactions and the psychological sense of support. In: Duck S, Silver R, editors. *Personal Relationship and Social Support*. London: Sage; 1990. p. 30–45.
21. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. *Psychol Bull*. 1985;98(2):310–57.
22. Roberts LJ, Luke DA, Rappaport J, Seidman E, Toro PA, Reischl TM. Charting uncharted terrain: a behavioral observation system for mutual help groups. *Am J Community Psychol*. 1991;19(5):715–37.
23. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J Mach Learn Res*. 2003;3:993–1022.
24. Nguyen T, Phung D, Dao B, Venkatesh S, Berk M. Affective and Content Analysis of Online Depression Communities. *Affect Comput IEEE Trans*. 2014; 5(3):217–26.
25. Dao B, Nguyen T, Phung D, Venkatesh S. Effect of Mood, Social Connectivity and Age in Online Depression Community via Topic and Linguistic Analysis. In: Benatallah B, Bestavros A, Manolopoulos Y, Vakali A, Zhang Y, editors. *Web Information Systems Engineering – WISE 2014*. Lecture Notes in Computer Science. 8786. Thessaloniki: Springer International Publishing; 2014. p. 398–407.
26. Mitchell M, Hollingshead K, Coppersmith G. Quantifying the Language of Schizophrenia in Social Media. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. Denver: Association for Computational Linguistics; 2015.
27. Schwartz AH, Eichstaedt J, Kern ML, Park G, Sap M, Stillwell D, et al. Towards Assessing Changes in Degree of Depression through Facebook. Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. Baltimore: Association for Computational Linguistics; 2014.
28. Zhang L, Huang X, Liu T, Li A, Chen Z, Zhu T. Using Linguistic Features to Estimate Suicide Probability of Chinese Microblog Users. In: Zu Q, Hu B, Gu N, Seng S, editors. *Human Centered Computing*. Lecture Notes in Computer Science. 8944. Switzerland: Springer International Publishing; 2015. p. 549–59.
29. McCallum A. MALLET: A Machine Learning for Language Toolkit 2002. Available from: <http://mallet.cs.umass.edu>. Accessed 2 Feb 2016.
30. Evans MS. A computational approach to qualitative analysis in large textual datasets. *PLoS One*. 2014;9(2):e87908.
31. Mimmo D, Wallach HM, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. Edinburgh: Association for Computational Linguistics; 2011.
32. AlSumait L, Barbara D, Gentile J, Domenico C. Topic significance ranking of lda generative models. In: Buntine W, Grobelnik M, Mladenic D, Shawe-Taylor J, editors. *Machine Learning and Knowledge Discovery in Databases*. Berlin: Springer; 2009. p. 67–82.
33. Holt-Lunstad J, Uchino BN. Social Support and Health. In: Glanz K, Rimer BK, Viswanath K, editors. *Health Behavior: Theory, Research, and Practice*. San Francisco: John Wiley and Sons; 2015.
34. Jacobson N, Greenley D. What Is Recovery? A Conceptual Model and Explication. *Psychiatr Serv*. 2001;52(4):482–5.
35. Barney LJ, Griffiths KM, Banfield M. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. *BMC Psychiatry*. 2011;11(88).
36. Griffiths KM, Reynolds J, Vassallo S. An Online, Moderated Peer-to-Peer Support Bulletin Board for Depression: User-Perceived Advantages and Disadvantages. *JMIR Mental Health*. 2015;2(2):e14.
37. Liu B. Sentiment Analysis and Subjectivity. In: Indurkha N, Damerau FJ, editors. *Handbook of Natural Language Processing*. Boca Raton: Taylor and Francis; 2010.
38. Wang X, Zuo Z, Zhao K. The Evolution of User Roles in Online Health Communities – A Social Support Perspective. *Pacific Asia Conference on Information Systems*; 2015; Singapore.
39. Wang C, Kraut R, Levine J. Eliciting and Receiving Online Support: Using Computer-Aided Content Analysis to Examine the Dynamics of Online Social Support. *J Med Internet Res*. 2015;17(4):e99.
40. Wang X, Zhao K, Street N. Social Support and User Engagement in Online Health Communities. In: Zheng X, Zeng D, Chen H, Zhang Y, Xing C, Neill D, editors. *Smart Health*. Lecture Notes in Computer Science. 8549. Beijing: Springer International Publishing; 2014. p. 97–110.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



Chapter 7

Foreword

This chapter presents the last of five empirical studies included in this thesis and the last of four that investigated an aspect of participation in BlueBoard followed by a comparative analysis of differentially-engaged users.

Given that the focus of the previous empirical studies in the thesis was largely quantitative, the current study used a qualitative analysis to deepen our understanding of the nature of participation within the framework of the findings observed so far. In particular, since the previous findings regarding user characteristics were limited to the four outcomes collected at baseline (age, gender, location and user status) and that these characteristics were not found to be meaningful determinants of differences in user engagement, this study sought to investigate the characteristics of users as they were presented in their initial post in order to determine if there are other independent characteristics of users, or ways in which the users present themselves, that are predictive of users' subsequent levels of engagement.

The first phase of this study used inductive coding of content in 150 users' initial posts within a qualitative framework analysis to generate a formal classification system for differentiating between high- and low-engaged users based on the content in their initial post. The design of the classification system may have been influenced by personal biases of the investigator (myself), particularly given my exposure to the content analysis findings in Chapter 6. Therefore, the second phase of the study employed blinded ratings from two independent mental health researchers to investigate the validity of the formal classification system in predicting the engagement status of users from their initial post content.

Thus, the study sought to answer the following questions:

RQ 6a: How do BlueBoard users present themselves in their initial post?

RQ 6b: Are there systematic differences in the way that users with different levels of engagement present themselves?

Through the qualitative analysis, it was found that higher- and lower-engaged users differed in terms of their 'awareness' characteristics regarding interpersonal-, self- and mental-health-awareness. Consistent with Chapter 6, this difference in awareness resembles differences in conceptions of recovery. The diagnostic accuracy of the classification system was found to perform significantly better than chance at discriminating between the top 25 (top 1.5%) of users and others.

Note: This paper has been prepared for submission to *BMC Psychiatry* but has yet to be submitted.

Are first impressions predictive of future engagement?

Qualitative analysis and blinded judgements of users' initial post in a mental health Internet support group

Bradley Carron-Arthur, Marita Cooper, Dominique Kazan, Kathina Ali, Kathleen M Griffiths

Abstract

Background: Online peer-support groups are an increasingly popular source of health information and support. The way in which users participate in these groups varies substantially from one-time questions or comments to long-term engagement. Understanding who engages with these groups and how they engage may be important for optimal service delivery. In this study, we sought to determine whether the way in which active users present themselves in their initial post is predictive of their subsequent levels of engagement as reflected in the frequency of their posts.

Method: The study comprised three phases across two stages. Stage 1 involved a qualitative analysis of the nature of the content in the first posts of 150 users through an inductive content analysis followed by a framework analysis designed to inform the development of a formal classification system for differentiating between high- and low-engaged users. Stage 2a involved an exploratory assessment of the validity of the above classification system. During this phase, two mental health researchers, blinded to user level of engagement, independently classified each initial user post of the 25 highest-engaged users and 25 single-post users (total n=50) as being from either a 'high-' or 'low-' engaged user based on the Stage 1 classification system. Data were analysed using logistic regression with predictive accuracy assessed using ROC curves. Stage 2b involved a more stringent test of validity of the classification system using an extended sample that included an additional 75 users from across the spectrum of engagement (total n=125).

Results: The qualitative analysis identified seven broad content themes (health issue, group-referent statement, self-referent statement, treatment, circumstances, social support and providing social support) that did not directly differentiate between users. However, transcending these themes were overarching differences in user presentation style. Highly-

engaged users were more likely than other users to present themselves in a way that demonstrated 'high-awareness' characteristics (raised-consciousness levels regarding self-, interpersonal- and mental-health-awareness) in the content of their initial post (Stage 1). A classification system based on these findings performed significantly better than chance at discriminating the high-engaged users from low-engaged users (36 correct out of 50) (Stage 2a), and high-engaged users from among users across the full spectrum of engagement (77 correct out of 125) (Stage 2b).

Conclusions: The current study suggests that users who demonstrate higher awareness characteristics in their initial post may become more highly engaged in a mental health Internet support group. This finding may facilitate the early detection of users most likely to become highly-engaged and may prove useful for optimising mental health Internet support group community-building strategies. However, it also highlights the need to consider strategies that support low-awareness users who are likely to engage only briefly.

Introduction

Although the 12-month prevalence of common mental disorders is 20%, only 35% of people with such disorders access a mental healthcare service [1]. The use of online health resources has become more common with 28% of Internet users having accessed mental health information online [2]. Further, online peer-to-peer platforms are a popular source of health information and support [3]. In 2011, online health communities were found to have been viewed by 18% of all Internet users in the USA [4]. In 2013, it was found that 8% of all Internet users in the USA had actively participated in such a community by asking a question or responding to someone else [5]. This level of engagement with online peer-to-peer groups has increased substantially since 2003 when the prevalence of such engagement was reportedly only 1.5% [6]. While such communities are premised on a mutual exchange of information, it has been observed repeatedly that the majority of content in these communities is created by less than 1% of all members [7, 8]. Research is required to understand the nature of this phenomenon and the social dynamics within online peer-to-peer support groups in order to inform their best-practice community management and optimal service delivery.

To date, few studies have systematically investigated individual differences in the way that users participate in Mental Health Internet Support Groups (MHISGs) [9] or the factors which predict levels of engagement. Further, although 'posting frequency' is the most commonly used metric for characterising user participation in an online health community [9] and the metric is often assumed to reflect the value of a user to the community, there is little evidence to validate this assumption. To address these gaps in the literature, we have recently undertaken a series of studies, focused on the MHISG 'BlueBoard' to determine who uses the group [10], who communicates with whom [11], what they communicate about [12] and whether these factors are associated with usage levels [10-12]. These studies revealed that users were most commonly consumers, female, aged 20-34 years and from an urban area, but that engagement with and retention on the MHISG was not strongly predicted by these characteristics [10]. Nor, based on a modular model of the community structure, were the odds of user interactions with each other significantly predicted by commonality of their characteristics, such as age, gender, location, consumer/carer status or type of mental health concern [11]. However, the registration date of the most highly active users in each module was found to precede most of the users with whom they communicate [11]. Further, through a topic modelling algorithm, it was found that the content created by highly active users was consistent with the notion that they perform a role as active help providers to other users [12]. The most highly-engaged users also tended to write more content that was consistent with the consumer model of recovery, whereas other users' content appeared more medicalised and symptom-focused [12]. However, the modelling analysis was limited by its quantitative and automated nature and further content analysis using qualitative techniques is required to refine, clarify and extend our understanding of the nature of the differences between the higher- and lower-engaged users.

There is considerable heterogeneity in content across posts in an MHISG. However, all active users engage in the activity of writing their first post. This standardised context provides an opportunity to identify associations between a user's attributes and their subsequent level of engagement. Moreover, this context precedes the influence of subsequent variable factors, such as responses from other users. To our knowledge, no previous study has investigated whether the content of a participant's initial post is associated with subsequent behaviour on an MHISG. However, an investigation of this question has the potential to increase our understanding of MHISG dynamics. A greater understanding of the characteristics that differentiate users may also assist managers of peer-to-peer support groups to optimise their community building efforts.

Generally, online community managers aim to increase the number of members and activity in a community in order to increase its capacity and sustainability [13]. A recent paper on online health community development noted that facilitating user engagement is resource intensive, requiring a focus on building relationships, fostering a sense of belonging among users and rewarding helpful activity [14]. The authors argued that core members were 'vital' to the sustainability of the group due to their high engagement and influence on the community [14]. In this context, fostering the engagement of core members can be seen as a priority for community managers. According to the author, 'all new members are potential core members' [14]. However, the study did not systematically investigate the propensity of different users to become core members and failed to take into account the potential role of individual differences among members. If there are differences in users' initial presentations that predict future engagement, an understanding of these differences may aid in the development of tailored approaches to building communities which may be more cost-effective than generic interventions. Thus, understanding the nature of content in users' initial posts presents not only an opportunity to better understand the dynamics of an MHISG, but also an important practical opportunity to understand differences that may inform strategies for building community from the outset of active user participation.

The current study will investigate user initial posts in two stages. In the first stage, user initial posts will be analysed qualitatively in order to determine whether there are systematic differences in presentation between higher- and lower-engaged users and to formalise a framework that describes these differences. The second stage will investigate if the resulting framework predicts user engagement. This will be undertaken in two phases with the aid of independent ratings from coders blinded to user level of engagement. In the first phase, the validity of the framework will be tested through a preliminary evaluation of the ratings for the highest- and lowest-engaged users. It was determined a priori that if the findings of the preliminary phase prove positive, a more stringent study would be undertaken employing ratings of the posts of users from the full spectrum of engagement. We present the Method and Results of each of the above stages separately.

Stage 1: Qualitative analysis

Method

Data

Data were drawn from the log data of the Internet support group BlueBoard (blueboard.anu.edu.au) for the period 1 October 2008 to 23 May 2014, during which 131,004 posts were made by 2932 users. This dataset has been described previously in four studies [7, 10-12].

The Internet support group: BlueBoard

This service was provided by the Centre for Mental Health Research (previously The National Institute of Mental Health Research) at The Australian National University. BlueBoard comprised 10 forums in which users communicated about a range of mental health issues, primarily depression, bipolar disorder and generalised anxiety. The board was established in 2003 as a mood disorder group, closed in 2007 and 2008 for funding reasons and re-established on 1 October 2008 before finally closing on 30 June 2016. BlueBoard's second iteration, from which the current data are drawn, did not include the content or registrations of previous users. BlueBoard was designed with the aim of being used for research purposes, in addition to its overt function as a support group. All potential users were advised of the research function of the board prior to registration. Moderators did not actively participate in any of the forums. Rather, they enforced rules, for example, by editing posts to remove any personally identifying information, and alerting the infracting user via a private notification on the board. The sampled data for the current study contained eight instances in which the content of a post had been modified (e.g. name removed). For ethical reasons, the original content was not retained. This research was approved by the Australian National University Human Research Ethics Committee.

Sample

Of the 2,932 registered users who contributed at least one post, 57% (n=1,664) identified themselves as consumers, and 8% (n=238) as carers at the time of registration; the remainder specified 'other' or provided no response to the identifier question. Since the attributes of consumers and carers may differ, only users who self-identified as consumers during the registration process were included in the current study.

The sample consisted of 150 users from three levels of engagement: high-engaged (n=50), moderate-engaged (n=50) and low-engaged (n=50). The 50 members of the high-engaged group comprised all of the top consumer posters on the board, including 23 of the 29 users who comprised the top 1% of users by posting frequency, a group that has previously been designated in the literature as ‘super users’ [9]. The total number of posts contributed by these users ranged from 145 to 11,994. The 50 low-engaged users consisted of a random sample of the 634 users who had contributed only 1 post. The 50 users in the moderate-engaged group comprised a random sample of the 978 users whose total posts fell in-between the two other groups (2 to 144 posts). The selection process for the low and moderate-engaged users was undertaken by a colleague not involved in the data analysis using a Microsoft Excel random number generator.

The percentage of women in the high, moderate and low groups was 78%, 63% and 72% respectively, with one response not reported among the moderate group. The median age in these groups was 35-39 years (high), 30-34 years (moderate) and 25-29 years (low), with one response not reported among the high group. Most of the participants in each group made their first posts to the depression forum (see Table 1).

Table 1. Frequency of users in each forum by first post for each participant groups.

Forum	High (n=50)	Moderate (n=50)	Low (n=50)
Depression	23	25	31
Bipolar	14	8	5
Generalized Anxiety	5	4	2
Social Anxiety	0	9	2
Borderline personality disorder	2	3	3
Eating disorder	0	0	0
Panic disorder	1	0	0
Obsessive compulsive disorder	0	0	0
Carers	1	0	0
General discussion	4	1	7

Analysis

The data were analysed using a qualitative framework approach [15, 16] with the aim of developing a classification system for differentiating between high- and low-engaged users. The analysis was undertaken by BC with input from MC and KG. Following a process of familiarisation, a thematic framework was devised using the data-driven inductive approach to

thematic analysis, where data are examined for codes and patterns and themes are allowed to emerge from the data. The content of posts were indexed to the themes using Nvivo software and charted in a table organised by *level of engagement* in rows (ranked from highest to lowest total post frequency) and *content themes* in columns. These data were then inspected for differences between higher- and lower-engaged users with the aim of detecting and interpreting patterns that might reveal differences between users.

Results

There was substantial heterogeneity in the way users presented themselves in their initial posts independent of the level of user engagement. Initial posts varied in length from a four-word statement (“I’m on, who’s on?”) to a 1,737 word story detailing the user’s mental health history. The median was 183 words. Despite this, a number of broad content themes and content sub-themes emerged from the analysis of the topic content of the data (see Topics below). There was little evidence that these content themes directly differentiated between users of differing levels of engagement. However, further analysis of overarching patterns in the charted data identified presentation styles that transcended content themes and differentiated between higher- and lower-engaged users (see Presentation Style section below). The findings from each of these stages of the analysis are presented in turn below.

Topic themes

Themes: Seven broad content themes (health issue, group-referent statement, self-referent statement, treatment, circumstances, social support and providing social support) and 29 sub-themes emerged from the data (see Table 2).

The most prevalent theme in the data related to the nature of the user’s *health issue* (77% of all users). Within this theme, statements regarding the *nature of the mental health issue* was the most common sub-theme (66% of all users). This comprised statements regarding disorder type (e.g. “I have depression”) (43%) or descriptions of the symptoms that they were experiencing (e.g. “I’ve been feeling down”) (23%). Few users stated that their mental health issue had been diagnosed clinically (8%). Statements about the nature of the mental health issues frequently co-occurred with descriptions of their *severity* (34%). Sometimes severity was described clinically as either mild, moderate or severe (5%), but it was more common for other adjectives to be used (e.g. “terrible”) (16%). Statements about severity were often also

coupled with statements about how it affected quality of life or functionality (e.g. “my bipolar is in the more mild category really, I can hold down a full time job and most of the time you wouldn't know unless I told you”) (9%). Users frequently reported a timeframe in which the mental health issues had been affecting them (64%). Some users stated both a *long- and short-term timeframe* referring to total chronicity and the most recent episode respectively (e.g. “I have had depression for 10+ years. I am going through a rough patch at the moment”) (32%). The majority of these statements referred to a long-term timeframe that was in the order of years (57%); short-term time frames were commonly referred to as “recently” or “at the moment” (32%). Content related to co-morbid *physical health issues* e.g. diabetes, was relatively infrequent (8%).

The second most prevalent theme was *group-referent statement* (75%). This included any content in which the user related themselves or their post to the Board or other users within it. It was common for users to make opening remarks signposting their status as a *new member* in the group (31%) along with statements that either indicated that they were unsure about how to participate correctly (11%) or that implied that they had an understanding of the group because they had been reading posts on the board (7%). One quarter of users explicitly stated their *reason for participation* in the group (25%). Most of these users reported that they were seeking social support (emotional, informational or appraisal support) (17%), whereas other users referred to the board as an outlet through which they were seeking to ‘vent’ (5%). A *request for responses* (38% of users) in the initial post most commonly involved a generic request for help or advice at the end of a narrative of the user’s experience (22%). However, some users made specific requests for input (20%) most commonly in relation to medication (10%). *Other* group-referent statements included gratitude expressed proactively for help from other users (10%) and/or apologies for the post length or the possibility of unintentional misconduct (7%).

The third most common content theme was *self-referent statement* (61%). This comprised any statement which provided information about the user themselves (excluding information specific to the nature of their health issue). Statements regarding the sub-theme *coping (emotionally)* were made by 44% of users in their initial post. For some this took the form of reflections on how they were managing their mental health issue emotionally or even coping with the act of coping (e.g. “I’m not scared of being anxious anymore”) (15%). For others, this comprised statements regarding how overwhelmed they felt (e.g. “I’m finding things so hard right now”) (15%). The sub-theme *expertise (cognitive)* included content regarding beliefs or

knowledge about matters that affected the management of their health issue (34%). Similar to the *coping (emotionally)* sub-theme, *expertise (cognitive)* content generally either involved statements that were insightful e.g. “i get irrational thoughts- so now when i have a thought i have to go thru the checklist to ensure its a rational thought or not” (10%) or conversely statements that suggested a lack of coping (“i don’t know what to do!”) (11%). *Other* self-referent statements (15%), included content written by users about their other personal attributes, commonly regarding non-disordered aspects of their personality (10%).

The fourth most prevalent theme was *treatment* (57%). More users reported *negative experiences* (30%) than *positive experiences* (19%) with treatment. Commonly, positive experiences related to medication (13%), talking therapies (3%), and alternative approaches (e.g. spirituality) (3%); negative experiences referenced medication (15%), or a health professional (5%). *Attitudes and intentions* regarding treatment (25%) commonly included statements exhibiting fear or reluctance to seek professional help or to take medication (11%). Less frequently, users reflected on how these attitudes had changed after experiencing a treatment (4%). *Other* content about treatment was confined to statements of fact about the user’s treatment use (e.g. medication) (11%).

The fifth most prevalent theme was *circumstances* (53%). This included content regarding situational matters that affected or were affected by the user’s mental health issues. *Partner relationships* was the most common sub-theme (21%), and typically comprised content regarding the negative impact of a separation (9%), or stories of supportive or unsupportive relationships (5%). A range of further circumstances included *other relationships* (13%) that were either helpful or harmful, *livelihood* stressors (20%), impacts of the user’s mental health problem on *daily functioning* (9%), situations that caused *trauma & grief* (8%) and various *other circumstances* (18%).

The sixth most prevalent theme was *social support* (42%). This included any statement about social support excluding references to formal treatment. Within the sub-theme *shared understanding* (23%), many users expressed delight at the feeling of empowerment that accompanied their engagement with a community that understood the issues they faced (10%) (e.g. “i’ve been a little glued to reading the posts because finally i have found others who share my own struggles”). Other users lamented the feeling that other people did not understand them (7%) e.g. “Nobody understands me or what i’m going thru”. In the sub-theme of *stigma* (8%), users most commonly described personal experiences of stigma (3%) or the effect of

their perception of stigma (3%). *Other* content regarding social support (26%) commonly included statements about social isolation (10%).

The seventh and least-prevalent theme was *provision of social support* (19%). Most users initiated their own thread (63%) when making their first post. However, some users wrote their initial post in another user's thread and provided either *emotional* (7%), *informational* (7%), *appraisal* (11%) or *other* types of social support (6%).

Differences as a function of subsequent user engagement: Table 2 presents the percentage of users who contributed to each content theme and sub-theme as a function of subsequent level of engagement. Although there was some variability between groups in the frequency of content indexed under each sub-theme, there was no evidence of a systematic pattern that differentiated between the first-posts topics of those with differing levels of engagement.

Table 2. Frequency of content indexation

Theme	High-engaged users n/50	Moderate-engaged users n/50	Low-engaged users n/50	Total Unique Users n/150 (%)
Health Issue				116 (77%)
Nature of the mental health issue	32	35	32	99 (66%)
Long-term timeframe	31	26	28	96 (64%)
Short-term timeframe	15	19	14	48 (32%)
Severity	17	20	14	51 (34%)
Physical health issue	6	3	3	12 (8%)
Group-Referent Statement				112 (75%)
New member	17	14	16	47 (31%)
Reason for participation	13	15	10	38 (25%)
Request for responses	17	26	14	57 (38%)
Other	13	11	12	36 (24%)
Self-referent Statements				91 (61%)
Coping (emotionally)	26	21	19	66 (44%)
Expertise (cognitive)	16	18	17	51 (34%)
Other	7	5	11	23 (15%)

Treatment				86 (57%)
Positive experience	11	10	8	29 (19%)
Negative experience	15	16	14	45 (30%)
Attitudes and intentions	11	18	8	37 (25%)
Other	9	11	5	25 (17%)
Circumstances				80 (53%)
Partner relationship	12	11	9	32 (21%)
Other relationship	9	7	4	20 (13%)
Livelihood	10	10	10	30 (20%)
Trauma & grief	5	4	3	12 (8%)
Daily functioning	3	4	6	13 (9%)
Other	10	7	10	27 (18%)
Social Support				63 (42%)
Stigma	4	2	6	12 (8%)
Shared understanding	12	14	9	35 (23%)
Other	14	12	13	39 (26%)
Providing Social Support				29 (19%)
Emotional support	5	1	5	11 (7%)
Informational support	3	3	4	10 (7%)
Appraisal support	5	4	8	17 (11%)
Other	4	1	4	9 (6%)

Presentation style

Although, as noted above, the inductive content analysis failed to reveal any direct associations between topic themes or sub-themes and engagement levels, the broader, integrated framework analysis did identify some consistent differences in the *presentation style* of users associated with engagement level. These commonalities in style were not exclusive to the formalised groupings but they were most apparent at the extremes of the spectrum of engagement, particularly among the most-frequent posters in the high-engaged group. We focus here on reporting the overarching differences in pattern for the higher-versus lower-engaged users.

The key finding from the framework analysis was that at the time of their initial posts, higher- and lower-engaged users differed in their *level of awareness* about mental illness. This difference was observed in three domains: *interpersonal-awareness*, *mental-health-awareness*

(knowledge and experience) and *self-awareness*. The differences between higher- and lower-engaged users in each of these domains are illustrated by the quotations in Table 3. These 20 pairs of quotations were obtained from the first posts of 18 users in each of the higher and lower categories (n=36 total). Only quotations from the predefined high- and low-engaged users have been included in this table, with the exception of one quotation each from two lower-engaged users who contributed 2 and 3 posts in total. The majority of the high-engaged user quotations (n=12) were sourced from among the top 14 most highly-engaged users. Although in combination, the quotations in Table 3 broadly represent the prototypical higher- and lower-engaged user, no single user expressed themselves in all the ways depicted in the table. Rather, Table 3 highlights the pattern in expression that was typical of higher- and lower-engaged users.

Differences in awareness in the ‘interpersonal’, ‘mental health’, and ‘self’ domains are detailed in turn below.

Interpersonal-awareness

Differences between higher- and lower-engaged users in interpersonal-awareness were observed in posts relating to the content themes *Social Support* and *Group-referent Statements*; and are illustrated by Quotations 1-5. Social isolation and the need for social support were common concerns for many users. Users often explicitly stated that BlueBoard assisted in addressing those concerns. However, users differed in their understanding of what social support BlueBoard might provide for them and, conversely, what support they might provide others.

Higher-engaged users typically demonstrated greater awareness of the community they were entering. For example, in Quotation 1H, the user explicitly acknowledged the community by stating that they had already been reading other users’ posts. Similarly, quotations 2H to 5H describe what the users hoped to contribute and gain from participating in the community. These quotations suggest that the high-engaged users were aware of what engagement in the community entailed, including for example, shared understanding, empowerment and timely support. In addition, these high-engaged users expressed themselves in a manner that was explicitly courteous to other users. In contrast, lower-engaged users typically did not demonstrate this level of community awareness. The latter was often characterised by an absence of introductory or closing remarks that referred to other community members. Other evidence of this limited awareness is illustrated by quotation 1L, which suggests that the user

was less focused on the needs of other users and was primarily motivated by the need to address their own personal concerns. Moreover, as demonstrated by quotations 2L-5L, the lower-engaged users often lacked a clear understanding of the specific potential benefits of engaging in the community. Rather, they appeared to be motivated primarily by an urgent awareness that they needed help, but were unsure of the type of help the board could provide.

Mental-health-awareness (knowledge and experience)

Differences between higher- and lower-engaged users in mental-health-awareness were most apparent in posts relating to the content themes *Health Issues, Self-referent Statements and Treatment*; and are illustrated by Quotations 6- 15 and 3. Quotations 6-9 relate to the user's awareness of severity, chronicity and coping behaviour as it relates to their own and others' mental ill-health. At a qualitative level, the severity of a user's mental health issue did not of itself appear to be associated with level of engagement (although this may be limited by the subjective nature of the judgement involved). Rather, the *perspective* of higher- and lower-engaged users towards the severity of their mental health issues appeared to be an important differentiating factor. Relative to those who were less engaged, higher-engaged users presented as veterans of their mental health issues who were less distressed by the symptoms per se, but more troubled by the chronic or recurrent aspect of their condition. Quotations 6H and 7H illustrate that higher-engaged users were not only concerned with specific coping behaviour, but also with the sustainability of coping behaviour. Quotations 8L and 8H illustrate the contrast in perspectives of a lower-engaged user who is feeling overwhelmed by their mental health problems and a higher-engaged user who has developed a strategy for managing complex issues associated with their mental illness. Quotations 9H and 9L contrast circumstances which illustrate the higher-engaged user has greater experience in living with their condition.

Quotations 10-13 relate to health management strategies. Users' self-reported approaches for managing their health were an important differentiating factor between low- and high-engaged users. The clearest illustration of this difference can be seen in quotations 10H and 10L which contrast the proactive management of a high-engaged user with the relative sense of helplessness conveyed by the low-engaged user. Higher-engaged users described specific activities which indicated that they were proactive in managing their health, by for example, studying mental health literature (as illustrated by quotation 11H), keeping a medication journal (quotation 12H), and disclosing their mental illness in a job interview (quotation 13H).

In contrast, lower-engaged users typically did not have a thorough understanding of mental illness (quotation 11L), were not specific about their strategies for recovery (quotation 12L), and did not disclose their issues to others (quotation 13L). Higher-engaged users were not infallible in their selection and implementation of proactive strategies; however, they possessed the insight to reflect on their choices if the latter were not beneficial to their health.

Quotations 14-15 (and 3) illustrate the differences in perspectives in higher- and lower-engaged users towards recovery. Low-engaged users tended to adopt a binary perspective characterised by a belief that there is a specific solution that will trigger full return to a previous normal state of health (see quotation 14L). By contrast, the higher-engaged user in quotation 14H did not expect to recover in a way that was completely free from illness. Quotations 15H and 15L, subtly illustrate the difference between the two user types, with the lower-engaged participant framing recovery as light at the end of the tunnel suggesting the potential for a binary shift in health whereas the higher-engaged user frames success in terms of a 'manageable' albeit difficult process. In quotation 3H, the higher-engaged user wrote specifically about the potential role of BlueBoard in their recovery process with an awareness of how this role differed from that of a medical professional. In contrast, the lower-engaged user in quotation 3L conveyed a sense that they approached BlueBoard in the hope of a miraculous solution and were unsure of what help they needed.

Self-Awareness

Differences between higher- and lower-engaged users in self-awareness were most apparent in posts relating to the content themes *Self-referent Statements and Circumstances*; and are illustrated by Quotations 16-20. Higher-engaged users often expressed themselves using a reflective style that indicated a meta-awareness of themselves contextually in terms of affect, behaviour and cognition and how the latter related to their mental health issues. In contrast, lower-engaged users tended to write from the perspective of being in and overwhelmed by their circumstances and experiences. In quotations 16H and 16L, both users described the experience of deliberately isolating oneself. However, the lower-engaged user described isolation as a behaviour that is a function of their beliefs whereas the higher-engaged user described isolation as a feeling that was a function of their behaviour. The higher-engaged user demonstrated self-insight by reflecting on the link between their behaviour and their subsequent mental health issues, whereas the lower-engaged user was immersed in their feelings of loneliness. Similarly, the contrast between the reflective approach of the higher-engaged user and the sense of immersion experienced by the lower-engaged user is evident in

quotations 17-20. Here it can be seen that low-engaged users were self-critical for lacking a ‘sensible’ reason to be depressed (17L), believed there was no hope because they were unlike anyone else (19L), concealed their issues from people who could help (18L), and acted impulsively by suddenly ceasing medication (20L).

Table 2. Quotations illustrating differences typical of awareness of higher- and lower-engaged users.

Higher-engaged		Lower-engaged	
Interpersonal-awareness			
1H	Hi All, I've been reading posts for a while but got the courage to post this morning	1L	Hi..uh i dont know how this really works, but have finally decided i need to vent a little
2H	I would like to share my story with you, both as a cathartic experience for myself, and in the hopes that the lessons I have learned can help others before you walk the painful path that I have walked. And I am hopeful that others will share their thoughts so that I can continue to learn and improve myself .	2L	To be honest, I'm not sure what I am looking for as a response
3H	deep down I feel the need for understanding from you as you are all sitting on this roller coaster ride with me and this understanding that i seek is different to the one my pdoc has	3L	I almost feel like I am posting here hoping that someone will give me a million dollars and all my problems will go away. I know that's not going to happen, but it is a bit of a cry for help even though I know it's not coming.
4H	So, I'm just looking for support and also some hope that things will get better. Hugs to all of you! :)	4L	All im asking for is some advice is there anything i can do myself to ease the pain of everything?
5H	Looking forward to sharing experience with you all and hopefully giving and getting some support during the tough times!	5L	please help
Mental-health-awareness (knowledge and experience)			
6H	[I'm] just your usual hurting soul	6L	i'm struggling. i really am. i cant talk about how useless i feel to my friends.
7H	I feel sick with the journey of my mental health	7L	Over the last few months I have not been coping, spending most of my days crippled by high anxiety & depression
8H	lets just focus on one problem at a time.	8L	I guess I'm finding it hard to except that I might constantly have to deal with this indefinitely. How do you guys deal with the prospect of dealing with depression for the rest of your life?

9H	I was diagnosed with bi-polar 3 years ago and am on seroquel. I am stable and did hold down a full time job	9L	I've started to feel a lot depressed only a few months ago. I'm 19, no job,
10H	I'm in therapy and proactive in the management of my health and overall wellbeing	10L	I'm right in the thick of it at the moment and I don't know what to do!
11H	the way that I cope with my disorder is to study. I like to study mental disorders and this helps me a lot. I don't go to University I just study for my own enjoyment	11L	I am having trouble understanding the illness
12H	Had I not forgotten to take my medication 18 months ago, I probably would still be enjoying good mental health as I had for the past 10 years, however I did forget and it has been a long, hard road back. I now keep a daily diary and note the medication taken so that I have a record.	12L	they feel I should just get over it or go back on my meds. This is not an option for me any more. I am/was proud of my achievement but am hitting a wall at the moment and just need to work through it.
13H	I disclosed at my interview for the job and still got it!	13L	I still feel shy about it - I haven't told anyone at work, for example, and I probably should.
14H	[I] kinda tapped out on looking for the land of rainbows and lollipops	14L	How do I switch off this feeling?
15H	Whatever you do don't give up, it is hard but definatley manageable	15L	Keep positive and be strong because there is a light ad the end of the tunnel
Self-awareness			
16H	I shut myself away from everyone and then wander why i feel so alone	16L	I havent told anyone how I feel because there is no one real to talk to
17H	Depression is not sensible, logical, or emotionally balanced. It doesn't care how good you've got it. It flies in the face of everything that makes sense.	17L	i should be grateful, Im gorgeous, iv had a great life, ive travelled a little, im healthy physically, and have many things to look forward to, so why do i feel this way
18H	I often put on a happy face too, but have found outlets through a few other friends. Although sometimes I feel guilty for burdening them. Deep down, I know that they would rather I burden them than I suffer alone	18L	when I go to my therapist each week, you know how it is, false front, smiling everythings just fiiiine
19H	But then again I always thought to myself that everyone thinks they are different, don't they?	19L	I jsut want to talk to someone but im too fukin different i feel
20H	Don't stop suddenly, side affect can be worse than what you are experiencing now.	20L	I'd had enough, lost faith in my psychiatrist and told him where he could jam his pills

Stage 2: Validity testing

Validity testing of the Stage 1 findings employed quantitative methods in a two phase approach (a. Exploratory; and b. Challenge) as recommended for the preliminary stages of evaluating a diagnostic test [17]. The Method and Results for each of these phases will be presented separately in turn.

a. Exploratory

The exploratory stage involved a preliminary examination of how well independent raters could classify high- and low-engaged users on the basis of the awareness typologies that emerged from the Stage 1 qualitative analysis. We sought to determine whether these ratings could be used in a predictive model to differentiate the most highly engaged users from the least engaged.

Method

Raters

Two mental health researchers (KA and DK), neither of whom had been involved in Stage 1, served as the study raters.

Data

Data comprised the initial posts of the 25 most highly-engaged users and 25 single-post users selected at random.

Procedure

The raters were provided with a written summary of the classification system for higher- and lower-engaged users that emerged from Stage 1 of the study (see box 1 below). This summary described the three domains in which the groups differed.

Each rater was asked to independently classify each of the 50 initial user posts as being from a 'high' or 'low' engaged user by providing four ratings for each post. This included a binary rating for the presence of each of the three types of awareness and a score out of 10 to

indicate overall whether they thought the user was ‘high engaged’ (10), ‘low engaged’ (0) or not sure either way (5). Coders were blinded as to group membership and posts were presented in random order. Following their initial independent rating (for analysis of inter-coder reliability) the raters conferred to provide a consensus rating (for analysis of the validity of the classification system). Training on a practice set was not conducted due to the small number of available high-engaged users.

Box 1. Summary of the observed differences between high- and low-engaged users in the content of their initial post used to instruct the raters.

The difference between high- and low-engaged users is not discernible from the *nature* of the circumstances that they describe in their posts, but rather it is discernible from the *way* that they describe them.

In order to differentiate the users, focus on the perspective from which the user appears to view their issues. There are three main ways in which this perspective differs:

- Interpersonal-awareness: High-engaged users often show that they have sought to learn about the BlueBoard community, i.e. by reading posts, and considered how their first post may relate to others before writing it. They often appear to have a specific idea of why they want to engage with the community e.g. for empowerment, shared understanding and timely support. The way they express themselves also appears more courteous of other users. In contrast, lower-engaged users often do not demonstrate this level of interpersonal-awareness. They often appear to be making a generic request for help. They appear to be motivated primarily by an urgent awareness that they need help from other people, but are unsure as to what that help specifically is.
- Mental-health-awareness (knowledge and experience): High-engaged users tend to have a multifaceted appreciation of the various matters involved in managing one’s mental health proactively and the ongoing effort involved. In contrast, low-engaged users often express themselves in ways that reveal they think or hope something can make their mental health issues go away. This distinction extends to users who are providing advice to others.
- Self-awareness: High-engaged users often describe themselves and their experiences using a reflective style that is indicative of meta-awareness of self. This awareness involves recognition of helpful and harmful factors in one’s own emotional, behavioural and cognitive response to situations or experiences. In contrast, low-engaged users tend to write from the perspective of being in and overwhelmed by situations and experiences. This difference can be the most difficult of the three to discern. Don’t expect a high-engaged user not to say they’re overwhelmed. To the contrary, if a user is explicitly acknowledging that they are overwhelmed, it is indicative that they somewhat self-aware. The difference between the two types of users is only apparent from the style of writing that indicates the perspective from which they are viewing the distressful experiences.

Analyses:

Inter-coder reliability was analysed using Cohen's Kappa for the binary ratings in each of the awareness domains and Pearson's rho for the overall score out of 10. Logistic regression was used to develop a predictive model for the top 25 users by posting frequency from the binary and continuous awareness ratings. The relative goodness of fit for models was assessed through the Akaike Information Criterion, and the predictive power of the final model was determined by computing the area under the Receiving Operator Curve in SPSS version 22.0.

Results

The inter-coder reliability was low to moderate among self- (62% agreement, $k = .28$), interpersonal- (70%, $k = .42$) and mental-health- (70%, $k = .42$) awareness ratings. Correlation of the overall Score out of 10 was moderate ($r = .58$, $p < .001$). There were 16/25 high interpersonal ratings among the high-engaged group compared with 6/25 in low-engaged group. For mental-health-awareness, there were 11/25 in the high group and 7/25 in the low group. For self-awareness, there were 12/25 in the high group and 8/25 in the low group. The overall score out of 10 was normally distributed (Mean = 5.0 SD = .30); 22 users scored higher than 5.

Seven simple logistic regressions were undertaken separately for each of the three binary awareness scores; the overall score out of 10; and three binary scores indicating whether or not a user had received at least one, two or three awareness ratings of any kind (e.g. coded 1 for greater than or equal to 2 awareness ratings of any kind, and coded 0 for less than 2). Interpersonal awareness (OR= 5.63, $p < .05$), the overall score out of 10 (OR=1.34, $p < .05$) and the binary score for at least one awareness rating (OR=6.73, $p < .01$) emerged as significant. These factors could not be combined in a more complex model due to multicollinearity. The model with the best fit was the binary awareness score indicating that a user had received at least one high awareness rating (AIC=61.22). The diagnostic accuracy of this binary awareness rating in predicting user engagement was found to be significantly better than chance (AUC = .72, 95% CI = .58 - .87) as shown in Figure 1. This criterion produced 36 out of 50 correct classifications (sensitivity = .76; specificity = .68). It is likely that the blind coders achieved close to optimal scores for this data, given that a non-blind coding conducted by the first author (BC), achieved 38 correct classifications out of 50 and that, as acknowledged in Stage 1, the

classification system reflected a trend rather than an invariable difference in responses as a function of engagement.

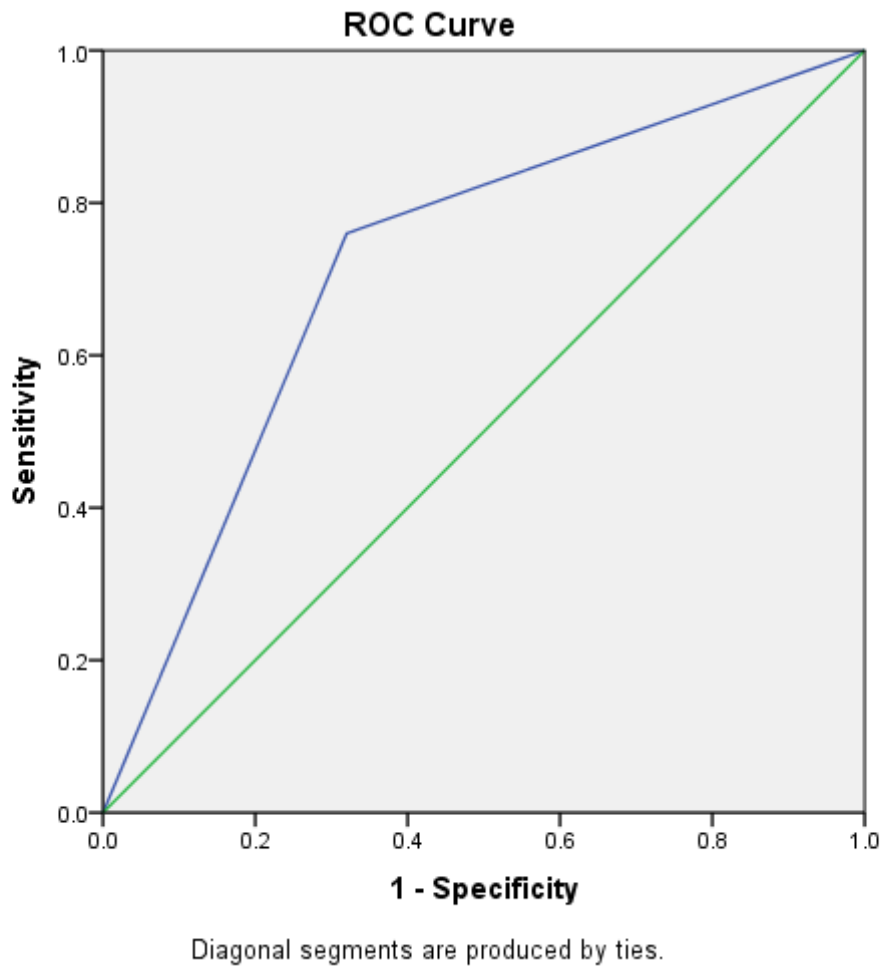


Figure 1. ROC curve showing the accuracy of the aggregate awareness rating (≥ 1 out of 3) for the classification of high- and low-engaged users.

b. Challenge

Since Stage 2a provided preliminary evidence of the validity of the classification system, we progressed to Stage 2b which involved a 'challenge' of the system employing the same methods as Stage 2a but with the addition of less clearly differentiated cases in the analysis. Again, we sought to determine if awareness ratings could be used in a predictive model designed to differentiate the most highly engaged users from among users across the entire spectrum of engagement.

Method

Data

In addition to the 50 post ratings from phase 1, data for this analysis included further ratings by the coders of the posts of the moderate-engaged group ($n=50$) and the lower-posting half ($n=25$) of the high-engaged users (in total, $n = 125$).

Raters and procedure

The previous raters (KA and DK) coded the additional posts following the same procedure as in the exploratory phase. They had not received feedback on their exploratory phase ratings.

Analysis

Inter-coder reliability for the additional 75 posts was analysed using Cohen's Kappa for the binary ratings in each of the awareness domains and Pearson's rho for the overall score out of 10. Logistic regression was used to develop a predictive model for the top 25 users by posting frequency from the binary and continuous awareness ratings. The relative goodness of fit for models was assessed through the Akaike Information Criterion, and the predictive power of the final model was determined by computing the area under the Receiving Operator Curve in SPSS version 22.0.

Results

The inter-coder reliability was low to moderate among self- (72% agreement, $k = .28$), interpersonal- (65%, $k = .31$) and mental-health- (67%, $k = .28$) awareness ratings. There were

28/75 high interpersonal-awareness ratings, 21/75 high mental-health-awareness ratings and 20/75 high self-awareness ratings. Table 3, which groups the sample into five equal groups of 25 according to the user's total post frequency, shows that higher awareness ratings were greatest among the 25 highest-posting users. The overall score out of 10 for the sample (n=125) was normally distributed (Mean = 4.6, SD = 1.8); 48 users scored higher than 5.

Seven simple logistic regressions were undertaken separately for each of the three binary awareness scores; the overall score out of 10; and three binary scores indicating whether or not a user had received at least one, two or three awareness ratings of any kind (as described in a) above). Interpersonal awareness (OR= 3.45, $p < .01$), the overall score out of 10 (OR=1.45, $p < .01$) and the binary score for at least one awareness rating (OR=4.37, $p < .01$) emerged as significant. These factors could not be combined in a more complex model due to multicollinearity. The model with the best fit was the aggregate binary awareness score indicating that a user had received at least one high awareness rating (AIC=117.50). The diagnostic accuracy of this aggregate binary awareness rating in predicting user engagement was found to be significantly better than chance (AUC = .67, 95% CI = .56 - .78). This criterion produced 77 out of 125 correct classifications (sensitivity = .76; specificity = .42).

Table 3. Frequency of 'High' rating for each theme of awareness across the spectrum of engagement grouped in bins of 25 users each.

Users (n=125) Posts Range	Interpersonal-Awareness (out of 25)	Mental-Health-Awareness (out of 25)	Self-Awareness (out of 25)
639 – 11,994	16	11	12
145 – 568	9	8	8
4-70	11	8	6
2 -4	8	5	6
1	6	7	8

Discussion

Principal Findings

To our knowledge, this study comprises the first systematic investigation of individual differences in the initial post of users of a mental health Internet support group and is the first study of the association between the content of the initial post and level of subsequent engagement. Seven topic themes and 29 topic sub-themes emerged from a thematic analysis of the first posts but these did not differentiate between high- and low-engaged users. Rather, a qualitative framework analysis suggested that higher-engaged users were more likely than lower-engaged users to demonstrate awareness of self, mental health and interpersonal matters, a finding confirmed in a quantitative study by the blind ratings of independent judges.

The concept of awareness

Since the term 'awareness' has been used in the extant black and grey literature on mental illness to convey different meanings in different studies and contexts [18-20] it is helpful to consider the concept as it emerged in the current study before reflecting on the key findings.

The concept of awareness seen in the current data was consistent with that of Andresen et al.'s [21] stage model of recovery from serious mental illness. In the latter model, the stage of 'awareness' refers to the individual's awareness of the personal agency required for recovery. In the current study, this was particularly apparent in the significant association between interpersonal-awareness and level of engagement, whereby the high-engaged users had a clear understanding of why they were participating (i.e. shared understanding, empowerment and timely support) and were proactive about attaining it, as opposed to low-awareness users who were uncertain about their needs and often made generic requests for help.

More broadly, the type of awareness reported in the current study is consistent with the consumer model of recovery, in which autonomy [22] and empowerment [23] are central components. Rather than focusing on cures that eliminate symptoms, as is the case in the medical model, the consumer model frames recovery as being 'recovery in' rather than 'recovery from' an illness [24]. It has been argued that awareness of this difference brings about a shift towards engagement in activities that support wellbeing, personal fulfilment and social connection and that complement medical treatments [23]. It is conceivable that understanding the recovery model requires a higher level of awareness of self, mental health

and interpersonal factors than is required for an understanding the less complex medical (curative) model.

Association between awareness and engagement

To our knowledge, the finding that initial awareness was greater among those who subsequently engaged more on the MHISG is novel, not previously having been reported in the literature. The finding has practical implications and raises questions about existing models of MHISGs (see sections below). In addition, the data increase our understanding of highly-engaged users of MHISGs.

In the literature, the top 1% of users by posting frequency have previously been categorised by some researchers as ‘super users’ [8]. However, the cut-off point between super users and others, and indeed the concept of ‘super users’ has lacked a meaningful definition beyond the operational definition of high-posting frequency [9]. Our finding that high-awareness characteristics were significantly more prevalent among the top 25 consumers by posting frequency (the top 1.5%) provides preliminary evidence that there is at least one meaningful factor associated with high posting frequency. The finding also reinforces our previous report from an automated topic model analysis of BlueBoard posts [12] that high- and low-engaged users may differ in their beliefs about models of recovery. Together this and the current findings suggest that low-engaged users may operate from the perspective of a medical model and that high-engaged users adhere to a consumer-recovery model.

With respect to the validity analyses of the formal classification system, sensitivity was adequate but the specificity was low indicating that some people demonstrating high awareness did not remain engaged with the board. It remains to be determined if the awareness measure captures factors that increase the *potential* for a user to become highly engaged. If so, the ‘false positive’ group may constitute users who could be successfully targeted with a suitable engagement promotion strategy, a possibility that merits careful investigation.

Models of MHISGs

The finding that subsequent engagement level can be predicted by the nature of the first post is not consistent with the ‘membership life-cycle’ model of depression Internet support groups recently proposed by Nimrod [25]. The latter model posits that users transition through a series of universal stages during their membership of a depression ISG. According to this model

all users are initially 'distressed newcomers' who subsequently transition to 'active help receivers' or 'passive followers', followed by 'relieved survivors' who either leave the group or become 'active help providers' for other users. The model was based on a cross-sectional survey of the users of 16 depression ISGs which found significant correlations between longer membership duration and stronger endorsement of statements about the benefits of participation ('I cope with the depression better', 'I am more capable in dealing with daily tasks' and 'I can be of help to others'). Implicit in this 'life-cycle' model is the assumption that the primary role of depression ISGs is to reduce symptoms and that the ongoing participation of users is motivated by the principal of reciprocity. The model does not take into account the possibility that people who are currently symptomless might participate in the group to help prevent a relapse. Nor does the model entertain the possibility that the focus of some users might be to optimise their recovery by pursuing a meaningful and purposeful life rather than to 'cure' their symptoms. Thus, although the life-cycle model may be applicable to some users - for example to those classified as low awareness at the time of first joining the MHISG - it is unlikely to be applicable to users with high-awareness at the time of joining.

An alternate view of the life-cycle of an MHISG involves a central group of high-awareness users who are engaged intimately in mutual support and who sustain the group. In parallel, low-awareness users intermittently approach the group for informal help or information and constitute a transient subset of the ISG membership which receives support from the high-awareness users. This model is consistent with the current findings, and those reported in our previous studies in which we identified that the highest-engaged users have high network centrality and that they registered significantly earlier than most other users in each of five generation-like cohorts in the BlueBoard community structure [11]. We also observed that they engage more frequently in consumer-recovery focused topics, whereas the other users engage in topics typical of a medical perspective of recovery [12].

Practical Implications

From a community management perspective, these findings have the potential to assist in the early identification of users who have the propensity to become highly engaged. The findings may be useful in increasing the efficiency of community building strategies for MHISGs. They also raise concerns about the appropriateness of attempts to encourage low-awareness users to become more highly engaged. The association between awareness and engagement may reflect the explicit intentions and needs of users. It cannot be assumed that encouraging further engagement of users, particularly low-awareness users, is necessarily in their best

interests. A recent study of an online breast cancer support group reported that users who sought and/or were provided with informational support were more likely to exit the group [26] whereas users who sought and/or received emotional support were more likely to remain engaged. This does not necessarily indicate that a good strategy for building a community involves providing all users with emotional support. Rather, the needs of low-engaged users might be best met in a different way to those of higher-engaged users given the brief period of their engagement. For example, it may be possible to trial messaging that refers users to formal sources of help or evidence-based treatment and other information, possibly tailored to the individual post using automated or non-automated methods. From both a research and clinical perspective it would be possible to track whether or not the initial user followed the link in the response to the user's initial post. The important future research question arising from these findings is therefore not simply – 'what community management strategies are beneficial generally?', but rather, 'what community management strategies are beneficial for what types of users?'

Limitations and future research

The current study aimed to investigate early predictors of high and low engagement on an MHISG. Ideally, this process would have involved a qualitative investigation based on data from a subset of the board, a developmental data subset to enable training on the classification system and a subset to test the classifier. However, due to the Zipfian nature of the frequency of posts on a bulletin board [7], and the findings of the qualitative analysis indicating that the concentration of high-awareness characteristics was only clearly visible in a small number (n=25) of highly-engaged users, there were insufficient high-engaged users to divide the board into data subsets. Accordingly, the current investigation serves as a preliminary exploration of awareness as a factor in subsequent ISG engagement. It provides a basis from which to undertake further research on similar ISGs to test the validity and generalizability of the current findings.

Given the subjective nature of the framework analysis, the blinded judgements were a potentially useful device for evaluating the validity and reliability of the framework. However, the inter-coder reliability was low, even after the initial set of coder practice, limiting the test's potential replicability. The low reliability may indicate a lack of clarity and specificity in the framework description in the rater guide or it may reflect a level of ambiguity in the posts. With respect to the former, descriptions were kept general without explicit examples to avoid inflating the accuracy of the test by overfitting the test to the current data. The broad nature

of the description does increase the potential generalisability of the findings to and testing of the framework in other contexts during which process the coding scheme can be refined.

Awareness characteristics only accounted for a portion of the variance associated with subsequent user engagement, and may not be a causal factor. A post hoc quantitative analysis indicated that high-engaged users (n=14) were significantly more likely than low-engaged users (n=5) to create their first post in the Bipolar Disorder forum, chi squared = 4.0 p< .05. It is possible that there is a separate moderator such as chronic illness or elevated levels of output during manic episodes which explains the observed findings. Standardised measures of illness type, severity and chronicity would enable this possibility to be tested more reliably than in the form it was available through the current study. Beyond user characteristics, interactions between users may be another major driver of user engagement [27]. This effect remains to be investigated in an MHISG.

Conclusion

The degree of awareness with respect to self, mental health and interpersonal matters demonstrated by a user in their initial post predicts the frequency of the user's subsequent engagement. High awareness in these domains is significantly more prevalent among the highest-contributing users (top 1.5%). This information may be useful for the early detection of people who could be supported to become core contributing members of the community. It also highlights that optimal community management must extend beyond facilitating engagement in the community to considering how the needs of low-awareness users may be best met in the brief period in which they are likely to be engaged.

References

1. Australian Bureau of Statistics. National Survey of Mental Health and Wellbeing: Summary of Results, 2007. Canberra, ACT: Australian Bureau of Statistics, 2008.
2. Fox S, Jones S. The social life of health information. Washington, D.C.: PewResearchCenter, 2009.

3. Griffiths F, Cave J, Boardman F, Ren J, Pawlikowska T, Ball R, et al. Social networks—the future for health care delivery. *Social Science & Medicine*. 2012; 75(12): 2233-41.
4. Fox S, Duggan M. Health Online 2013. Washington, D.C.: PewResearchCentre, 2013.
5. Fox S. Peer-to-peer Health Care. Washington, D.C.: PewResearchCentre, 2011.
6. Owen JE, Boxley L, Goldstein MS, Lee JH, Breen N, Rowland JH. Use of health - related online support groups: population data from the California health interview survey complementary and alternative medicine study. *Journal of Computer - Mediated Communication*. 2010; 15(3): 427-46.
7. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions*. 2014; 1(4): 165-8.
8. van Mierlo T. The 1% rule in four digital health social networks: an observational study. *Journal of Medical Internet Research*. 2014; 16(2): e33.
9. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From help-seekers to influential users: A systematic review of participation styles in online health communities. *Journal of Medical Internet Research*. 2015; 17(12): e271.
10. Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A. User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study. *Internet Interventions*. 2017; 7: 9-15.
11. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM. Community structure of a mental health Internet support group: modularity of user thread thread participation. *JMIR Mental Health*. 2016; 3(2): e20.
12. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Griffiths KM. What's all the talk about? Topic modelling in a mental health Internet support group. *BMC Psychiatry*. 2016; 16(1): 367.
13. Iriberry A, Leroy G. A life-cycle perspective on online community success. *ACM Computing Surveys (CSUR)*. 2009; 41(2): 11.
14. Young C. Community management that works: how to build and sustain a thriving online health community. *Journal of Medical Internet Research*. 2013; 15(6): e119.
15. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*. 2013; 13: 117.
16. Ritchie J, Spencer L. Qualitative data analysis for applied social policy research. In: Huberman AM, Miles MB, editors. *The Qualitative Researcher's Companion, Classic and Contemporary Readings*. Thousand Oaks, CA: Sage; 2002. p. 305-29.
17. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240(4857): 1285.
18. World Federation for Mental Health. World Mental Health Day 2017 [16/06/2017]. Archived at <http://www.webcitation.org/6rFgO5rWz>]. Available from: <https://1010.org.au/>.
19. World Health Organization. *The World Health Report 2001: Mental health: new understanding, new hope*: World Health Organization; 2001.
20. Wright A, McGorry PD, Harris MG, Jorm AF, Pennell K. Development and evaluation of a youth mental health community awareness campaign—The Compass Strategy. *BMC Public Health*. 2006; 6(1): 215.
21. Andresen R, Oades L, Caputi P. The experience of recovery from schizophrenia: towards an empirically validated stage model. *Australian and New Zealand Journal of Psychiatry*. 2003; 37(5): 586-94.
22. Australian Government Department of Health. *National Standards for Mental Health Services*. Canberra: Australian Government Department of Health, 2010.
23. Jacobson N, Greenley D. What is recovery? A conceptual model and explication. *Psychiatric Services*. 2001; 52(4): 482-5.

24. Davidson L, Roe D. Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health*. 2007; 16(4): 459-70.
25. Nimrod G. The membership life cycle in online support groups. *International Journal of Communication*. 2012; 6: 23.
26. Wang X, Zhao K, Street N. Analyzing and predicting user participations in online health communities: A social support perspective. *Journal of Medical Internet Research*. 2017; 19(4): e130. 28438725
27. Wang X, Zuo Z, Zhao K. The evolution of user roles in online health communities: A social support perspective. Pacific Asia Conference on Information Systems; 2015.

Chapter 8

This chapter commences with a summary of the key findings of this thesis (Section 8.1) and consideration of the research and its methods in the context of the extant literature (Section 8.2). A synthesis of findings is then presented in the form of a model of participation in an MHISG (Section 8.3) followed by a discussion of potential implications of this model in practice (Section 8.4); the strengths and limitations of the methods employed in this thesis (Section 8.5); possible directions for future research (Section 8.6); and a conclusion to the thesis (Section 8.7).

8.1 Summary of research findings

8.1.1 In what ways has participation previously been defined and measured in an MHISG, and other online health communities? (RQ1)

The systematic review conducted at the outset of the thesis found that many different metrics have been employed to characterise active participation in health ISGs (Chapter 2 [1]). The single most commonly employed metric of participation across all health ISGs, as well as MHISGs alone, was posting frequency, with higher posting frequency often viewed as an indicator of the value a user contributed to the community. A minority of the studies identified in the review combined metrics to create sophisticated measures of multi-dimensional, specific participation styles, such as ‘influential users’ or ‘key players’. However, there was little overlap in the multi-dimensional measures employed across studies and little indication of the convergent or divergent validity of the measures.

Many of the participation styles identified have not yet been investigated in an MHISG and it may be beneficial to do so in future. However, it was clear from the review that there remained major unanswered questions relating to MHISG user engagement as measured by the single metric posting frequency. It was, for example, not acknowledged by the literature that differences in posting frequency may reflect *qualitatively* different *styles* of participation in an MHISG. Rather, posting frequency has typically been regarded in the literature as a unidimensional measure of the level of engagement in a single type of activity in which all

users participate. Given the extent to which these assumptions regarding posting frequency and engagement were embedded in the literature and the other extant gaps in research on this common metric, we regarded this topic as a priority for research.

8.1.2 What is the nature of active participation in BlueBoard (in terms of user characteristics, activity, connectivity and content) and does it differ between differentially-engaged users? (RQs 2-6)

The empirical studies in this thesis investigated user characteristics, connectivity and content to determine if these attributes were associated with differences in posting frequency. Some differences in user characteristics were apparent, including higher posting frequency by consumers compared to carers and others, fewer posts for those younger than 20 years compared to the 20 to 34 and 35 to 50-year-old groups, and a trend towards lower posts among rural and remote residents compared with their urban counterparts (Chapter 4 [2]). However, the predictive value of these demographical and clinical characteristics was very small (Chapter 4 [2]).

More striking was the observation that highly-engaged users presented differently to lower-engaged users in their initial post (Chapter 7 [3]). Higher-engaged users tended to demonstrate raised consciousness levels in multiple 'awareness' domains, including awareness of self in terms of affect, behaviour and cognition; awareness of mental health with respect to knowledge and experience; and, particularly, interpersonal awareness relating to the benefits to self and others of engaging with an MHISG (Chapter 7 [3]). Furthermore, this difference was consistent with the findings from a large-scale automated quantitative study on all post content which revealed systematic differences in the topics on which higher- compared to lower-engaged users focused (Chapter 6 [4]). The latter study concluded that higher-engaged users posted on topics about recovery in mental illness that aligned with the consumer-based model of recovery whereas their lower-engaged counterparts contributed on topics consistent with the medical model (Chapter 6 [4]); the qualitative findings from the study in Chapter 7 were consistent with this interpretation.

The position of the most highly-engaged users in the network was also found to differ systematically from that of the lower-engaged users (Chapter 5 [5]). Rather than there being a modular community structure that revolved around homophily in user characteristics or that was delineated by structural boundaries set out by the different topical sub-forums, high-

engaged users were found to be central to five groups that could be interpreted as generational cohorts across the almost six-year lifespan of the MHISG (Chapter 5 [5]). The studies do not provide evidence of causality. However, given that highly-engaged users registered significantly earlier than the median user in each cohort (Chapter 5 [5]) and that there were systematic differences in the way they presented prior to any other interactions with users (chapter 7 [3]), it is possible that the social dynamics of the group were markedly affected by the highly-engaged users. Consistent with this possibility, it was observed that highly-engaged users adjusted the content of their posts in response to the topics initiated by lower-engaged users (Chapter 6 [4]). This suggests that highly-engaged users perform an active help provider role, albeit as a minority component of their overall activity. The majority of activity (posts) by higher-engaged users took place in the threads of other highly-engaged users. This suggests that they were also performing a role as help providers to other highly-engaged users. Significantly, the types of content used in responses to other highly-engaged users differed significantly from that of the content in the responses to low-engaged users, with a greater use of topics aligning with consumer rather than medical conceptions of recovery. This suggests that there may have been a difference in the types of help provided by the highly-engaged users to these two groups.

8.2 Findings and methodology in the context of other research

The empirical studies in this thesis have contributed several novel findings to the field of MHISG research using methods and analyses not previously applied to this topic.

Firstly, a power law (Zipfian distribution) had not previously been fitted to or reported to describe the distribution of participation in an MHISG (Chapter 3 [6]). However, the findings are consistent with those previously reported for a range of other digital and non-digital phenomena [6], providing further evidence of the ubiquitous nature of the Zipfian distribution.

Secondly, user characteristics data collected during registration has not previously been reported in a study of an ISG for depression and related mental disorders to reduce the selection biases observed in survey data studies (Chapter 4 [2]), although such data has been reported in studies of problem drinking and smoking IGs [7, 8]. Accordingly, the validity of the resulting findings is likely to be greater than those from previous studies of the MHISGs for

conditions such as depression. Certainly, the findings from the current study indicate that respondents in some surveys are not representative of all registrants. For example, a recent study employing a cross-sectional survey [9] of respondents to an advertisement posted on 40 different MHISGs yielded a sample in which the average self-reported participant membership duration was four years. This figure is very high in comparison to BlueBoard, in which only 3.7% of members were active for more than a year, and 49.0% were active for less than a day [2], which together with the known distribution of activity on ISGs in general suggests that the survey sample was biased towards the very highest-engaged users across those MHISGs. This points to the need for a more careful consideration of research design in future studies of the characteristics of ISG users.

Thirdly, the analysis of user sub-group modularity in Chapter 5 [5] and the topic modelling technique implemented in Chapter 6 [4] each employed methods that were novel in the field of MHISGs and provided a new perspective on large-scale patterns of participation that have not previously been reported in the literature and which may otherwise not be visible to community managers.

Fourthly, no study had previously systematically investigated the way in which users first present to an MHISG, nor taken an inductive approach to content analysis to determine if there are systematic differences between differentially engaged users (Chapter 7 [3]). The study was the first to demonstrate differences in awareness characteristics between the high- and low-engaged MHISG users.

Overall, the mixed-methods approach employed across the thesis has enabled the triangulation of findings, particularly regarding potential differences in conceptions of recovery among differentially engaged users.

Since the publication of the three articles in Chapters 3, 4 and 5, subsequently published research has yielded results that are consistent with their findings; that is, the distribution of engagement is repeatedly and optimally described by a power law [10], user characteristics (demographics and indicated symptoms) are not strongly predictive of engagement [11], and the modularity in community structure is broadly cohesive and united across different topical threads by highly-engaged central users, [12]. The latter was observed, however, in an ISG that was dedicated to a single condition, problem drinking, a condition for which, in contrast to the Chapter 5 study, modularity cannot be differentiated by different condition-specific interests.

The empirical studies in Chapters 6 and 7 have not yet been investigated or replicated in another MHISG. A key finding of each of these studies was the differing conceptions of recovery that were apparent in the content of the posts of higher- and lower-engaged users. While the distinction between medical and consumer models of recovery may be particularly relevant to mental health, it is also consistent with a broader trend in online health communities in which higher-engaged users tend to engage more in emotional support and companionship than lower-engaged users who tend to focus more on informational support [13-15]. Indeed, this distinction may be consistent with an observation regarding prototypicality in online communities generally. A recent study using natural language processing techniques to analyse the nature of three different online communities found that the prototypicality of a user, as measured by the nature of their language use, was predictive of leadership status as judged by peers [16]. The notion that leaders in an MHISG are those who are most prototypical of the group is consistent with well-established social psychological theories of group membership [17]. This may be exemplified anecdotally by a report from a recent study in which participants who self-identified as having depression and/or anxiety were directed to participate in a long-established MHISG (Psych Central). One participant who encountered difficulties in engaging in the intervention reported: 'It seemed that the majority of the regular posters on Psych Central went way beyond a tad anxious or a bit blue. A lot of the members had severe mental illnesses or told stories about going through horrendously traumatic experiences. I felt a little over my head in the community.' [18].

Although it is not the subject of research in the current thesis, it is conceivable that the clinical outcomes of MHISG users with different participation styles differ. It is also possible that the current findings shed some light on the inconsistencies in findings in the current literature. As noted in Chapter 1, Griffiths et al [19] reported that a depression ISG intervention purpose-developed for the research study resulted in a significantly greater reduction in depression than an attention control condition. By contrast, Dean et al [18] found no difference in depression outcomes for participants referred to a pre-existing MHISG compared to participants encouraged to engage in an expressive writing exercise, and that the observed changes in outcomes were similar to those recorded in no-treatment control groups of other studies that employed computer-based treatment programs. In the past, Barak et al [20] has asserted that research on purpose-built ISGs lack ecological validity as they lack the therapeutic components that occur in 'real' ISGs. In this context, the differing findings from the two studies might appear somewhat paradoxical. However, in the absence of a pre-established culture and leadership, it is possible that the purpose-constructed MHISGs provided

participants who might otherwise have failed to engage or identify with a pre-existing MHISG, a greater opportunity to engage in mutual help [19, 21, 22]. Previously, studies have shown that there is an association between higher levels of activity in an MHISG and better outcomes in terms of self-reported emotional distress in the content of posts [23] and depressive symptoms as measured by a cross-sectional survey [24] and longitudinal surveys [25]. However, none of these studies employed a randomised controlled trial design. Thus, it is not clear if this is a causal relationship, nor, if it is causal, whether higher posting frequency is beneficial in of itself, or whether it is simply a proxy measure of different types of peer support that differ in their benefits. Further research could test these hypotheses.

Certainly, it is clear that there is a dearth of knowledge regarding the experience of low-engaged active help seekers of whom the majority of users in ‘real-world’ MHISGs are comprised. It is possible that MHISGs do not reduce symptoms for the majority of low-engaged users. However, Barak et al. have asserted that there may be other more pertinent potential benefits of MHISGs [20]. It has been suggested that in ‘real life’ the impetus for seeking help from an MHISG comprises a combination of symptoms, social isolation, stigma, convenience and reluctance to seek professional help, and that the benefit of engagement is not symptom reduction per se, but rather the resulting consumer activation, challenge to self-stigma and help seeking from a professional [26, 27]. There is evidence based on self-reported data that for a substantial minority of users (36%), the MHISG has been a catalyst for formal help-seeking, and that active users are more likely to have sought formal support than passive users [9], albeit that the latter may reflect pre-existing characteristics that prompted active participation in the MHISG. Further research is required to investigate the factors associated with professional help seeking in association in MHISG use, particularly among low-engaged users.

8.3 Synthesis: A model of participation

The current findings, together with previous research, provide a basis on which to propose a new model of MHISG participation. This model, referred to here as *The Tripartite Model of MHISG Participation*, comprises four styles of participation (*passive followers and help seekers, mutual helpers, active help seekers, and active help providers*) and is illustrated in figure 8.1. A

summary of the characteristics which define each participation style, arising from studies in this thesis, is provided in table 8.1.

The proposed model contains groupings of participation styles at three levels with transitions between the four styles of participation. It shares some attributes of the 'membership life-cycle' model [28] (see Chapter 1) in that both the latter and the Tripartite model include the 'passive' and 'active help-seeker' participation styles and transitions between roles. However, the Tripartite model also includes *mutual helpers*. The life-cycle aspect in this model is represented by the replications of the same structure which depict the generational cohorts observed in Chapter 5. This illustration has strong similarities with the graphical representation of the BlueBoard community structure in figure 2 of Chapter 5, which highlights the high-engaged users as central to the network surrounded by many low-engaged users in the periphery. This model also draws on suggestions that online communities are often characterised by a core-periphery structure [29, 30].

At the top layer of the Tripartite model are *mutual helpers*. High engagement (as measured by posting frequency), high awareness characteristics, network centrality, early registration relative to the peers with whom they communicate and a tendency to communicate about topics that resemble the consumer-defined notion of recovery are defining characteristics of this participation style.

The second layer of the Tripartite model comprises *active help seekers* and *active help providers*. In contrast to the *mutual helpers*, these two participation styles are characterised by low engagement, low awareness characteristics, low network centrality and a tendency to communicate about topics that resemble the medical-defined notion of recovery. *Active help seekers* and *active help providers* are distinguished from each other based on the direction of help being provided, as indicated by the content of posts or activities such as starting a thread in order to seek help, or responding in a thread in order to provide help to another user.

At the base of the Tripartite model, there are *passive followers and help seekers*. Previous research has defined this group as those who read but do not actively participate; they are often referred to as 'lurkers' [31].

The arrows in the diagram represent transitions between participation styles. Hypothetically, any user may transition between any two participation styles and back again. The diagram only

includes arrows for which there is empirical evidence of the transition. *Passive followers and help seekers* are believed to benefit vicariously [32, 33]; however, they have also been observed to transition to the active participation styles: *mutual helpers*, *active help seekers* and *active help providers* (Chapter 7 [3]). *Mutual helpers* benefit from each other, however, they often transition in their participation style temporarily to serve as *active help providers* for the *active help seekers* (as indicated in the diagram by the two-way arrows).

The difference between *active help providers* and *mutual helpers* is defined by the extent to which the interactions between users are transient and asymmetrical. *Active help providers* are engaged in relatively transient interactions with other users, whereas *mutual helpers* are engaged in more sustained and symmetrical relationships with each other (resulting in higher posting frequency). The nature of these types of peer support, in terms of the types outlined in Chapter 1, is akin to mutual help between the *mutual helpers* and a consumer-run service between the *active help providers* and *active help seekers*. In the case of the latter, it involves users with low levels of awareness seeking help from consumers who are more advanced in their recovery and who have higher levels of awareness [34] and are providing their expertise in a voluntary rather than professional capacity. Hypothetically, the *active help providers* may also provide support to *mutual helpers*, although this was not directly investigated by a study in this thesis and requires investigation. In contrast to the conclusions from previous cross-sectional survey research (and the corresponding Membership Lifecycle model) which has suggested that there is a gradual progression from the role of low-engaged distressed newcomer to high-engaged help provider [28], the current research suggests that high-engaged *mutual helpers* are also potentially differentiated from low-engaged active users from the outset of their participation by their awareness characteristics.

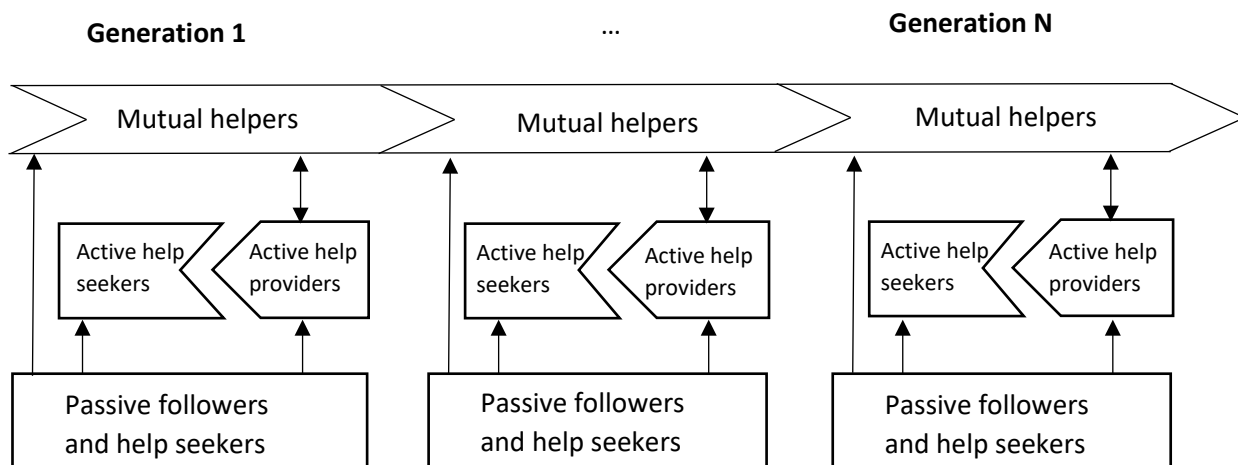


Figure 8.1. Tripartite Model of MHISG Participation. The arrows indicate transitions in participation style that have been evinced in the current thesis, although it is plausible that a user may transition between any two participation styles. Participation styles are outlined in table 8.1.

Table 8.1. Defining metrics of the participation styles in the Tripartite model

Participation style	Metrics
Mutual helpers	<p><u>Activity</u>: High posting frequency</p> <p><u>Characteristics</u>: High interpersonal awareness, consumer, older, registration early relative to connected users</p> <p><u>Connectivity</u>: high network centrality</p> <p><u>Content</u>: typical of consumer model of recovery</p>
Active help providers	<p><u>Activity</u>: variable but often same as mutual helper, responding in a thread</p> <p><u>Characteristics</u>: variable but often same as mutual helper</p> <p><u>Connectivity</u>: variable but often same as mutual helper</p> <p><u>Content</u>: typical of medical model of recovery</p>
Active help seekers	<p><u>Activity</u>: low posting frequency, starting a thread</p> <p><u>Characteristics</u>: low awareness</p> <p><u>Connectivity</u>: low centrality, peripheral in network</p> <p><u>Content</u>: typical of medical model of recovery</p>
Passive followers and help seekers	<p><u>Activity</u>: No posting frequency</p> <p><u>Characteristics</u>: No evidence from the current studies</p> <p><u>Connectivity</u>: Not applicable</p> <p><u>Content</u>: Not applicable</p>

The most important contribution of the proposed Tripartite model is the shift from a dualistic model, in which participation is conceptualised only as active or passive, to a more refined model in which active participation is further differentiated by type of peer support. In the Tripartite model, the MHISG is sustained by a core group of users who are engaged in long-term symmetrical relationships with one another with the aim of achieving personal recovery. Peripheral to this are the majority of users who approach the group and engage in transient asymmetrical interactions with the core members and one another. The core members are engaged in mutual help and the peripheral members engage briefly with a consumer-run service. Although posting frequency alone does not discriminate between these different styles of participation, the studies in this thesis demonstrate that frequency is associated with these styles and may therefore serve as a legitimate proxy measure for them.

8.4 Implications for practice and policy

The main implication of the current research for practice is that MHISG service delivery may be optimised by modifications that account for differences in user characteristics and participation styles. The empirical studies in this thesis were confined to descriptive and correlational analyses. We did not undertake studies of the effect of tailoring interventions on symptoms and recovery. Accordingly, it is not possible to recommend specific, evidence-based tailored interventions for users. However, the findings in the thesis provide a basis on which to suggest modifications to an MHISG that may be beneficial and that could be investigated by future research.

8.4.1 Presentation of the MHISG to new users

In Chapter 7, it was observed that low-engaged users often had low interpersonal awareness of the community to which they were presenting and low awareness of how participation in the community might be of benefit to them, as was illustrated by the user who wrote 'I don't know how this really works' (see p. 101). New users in an MHISG may therefore benefit from a description of the community, the different ways in which people participate and the different ways a person may benefit from participating in the community. This thesis has observed diversity in MHISG participation, with clear differences in the nature of the participation of users who have different posting frequencies. It cannot be assumed that encouraging more

participation from all users will be of benefit to them. It may be more beneficial to focus on the quality of that participation, regardless of the level or duration of the activity.

The research in this thesis has provided important insights into the nature of participation in an MHISG. However, the model of participation emerging from the thesis is currently not reflected in the representation of the community on home pages of MHISGs. This is illustrated by a perusal of the home pages of several well-established MHISGs including, for example, the SANE forums and Psych Central (see figures 8.2 and 8.3) that use a similar presentation style and layout to BlueBoard. By displaying only registration statistics, these home pages convey the impression that the community is comprised of a large number of people with conversations distributed across multiple different sub-forums and threads. Commonly, MHISG forums show that there are thousands of threads and hundreds of thousands of posts. This information is factual. However, standard statistics do not convey the information that the vast majority of activity on the boards is actually the product of a small group of members, and that although a community may appear to be distributed across multiple sub-forums, the community is actually coherently united by this core group of users. It is perhaps not surprising then that most users do not present with a clear understanding of the MHISG (Chapter 7).

Assuming the findings from the current thesis generalise to other MHISGs, there may be value in providing a description of the nature of the community and common participation styles within it to new users. In particular, as detailed in box 8.1, it may be helpful to clarify that although the community may seem large, the number of highly-engaged users is substantially smaller [6] and that these users are often responsive and willing to offer support to new users [4].

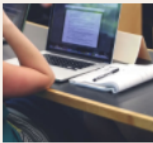
Box 8.1: Suggested additions to introductory messaging for new MHISG users

Explain that:

- (i)** Although there are many forums in which to participate in discussions, there is a cohesive and supportive community that transcends these [5].
- (ii)** The users who are most likely to respond to a new thread tend to participate in multiple sub-forums [2]. It is therefore not necessary to introduce yourself wherever the latest activity is occurring to obtain a response. Some MHISGs signal to new users that they should introduce themselves in a specific sub-forum [35].
- (iii)** It is common for people to present to an MHISG when they are distressed even though they are unsure how it will help them and unsure how to participate [3]. (This message aims to normalise all forms of participation and ensure some styles of participation are not discouraged implicitly or inadvertently.)
- (iv)** Forum users have different approaches to recovery, which may lead some members to participate more frequently than others [4]. Some highly involved members engage in ongoing participation to support their ongoing recovery [36], whereas others engage short term in order to address and resolve information needs [37] or receive timely support [36]. Users may also benefit from an explanation of the differences and similarities between consumer (personal) and medical models of recovery.

Consistent with current best practice in the development of service delivery models in mental health [38], users should be consulted and involved in the development of this messaging [39].

Overall, the aim of these suggested modifications to introductory messaging is to provide a transparent and unambiguous experience for new users. The messaging is not designed to encourage greater participation rates or activity, but rather to assist users to self-determine whether and how active participation may be best suited to their needs. However, as acknowledged above, the effectiveness of this approach compared to standard practice requires empirical evaluation.



Q&A Info Session: NDIS
Lived Experience
Ask us anything about the NDIS!

New here? Introduce yourself
Lived Experience

Tell us what you think of the Forums
Lived Experience

Discussion Topics

<p>Our experience & stories Living with a mental health issue can be lonely. Sharing stories helps us to find others who have faced similar experiences and to support each other. What's your story?</p>	<p>Posts 27582 Views 206580</p>
<p>Looking after our wellbeing Every day is a new day, with its achievements and its challenges. What helps to look after yourself and improve wellbeing?</p>	<p>Posts 19423 Views 68199</p>
<p>Enjoying time with others Whether we are sharing our favourite book or what we did yesterday, connecting with others can be good for us. How do you spend time with others? Join the conversation and enjoy some time with other members.</p>	<p>Posts 53849 Views 60537</p>
<p>Something's not right When something's not going well, it helps to share this with others and learn how they have coped and found solutions. Is there something you do that others would find helpful too?</p>	<p>Posts 36485 Views 133466</p>
<p>What's new: services, research, and technologies The things that can help us look after ourselves are constantly changing. Find out and share information about what's new. Have you heard about something you think would help others?</p>	<p>Posts 1308 Views 31261</p>

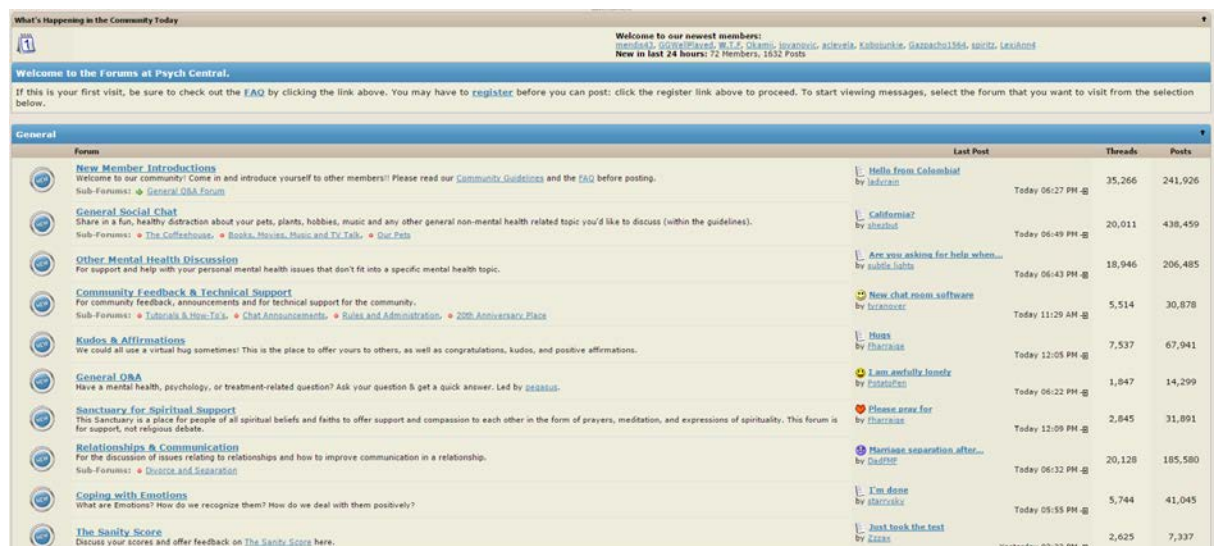
Forum Guides

- [Getting started](#)
- [Community Guidelines](#)

Latest Comments

- [Re: Night Shift](#)
- [Re: Am Not Coping](#)
- [Re: Am Not Coping](#)
- [Re: General chit chat](#)
- [Re: Night Shift](#)

Figure 8.2. SANE Forums Home Page



The screenshot shows the Psych Central Forums Home Page. At the top, there's a navigation bar with a search icon and a welcome message: "Welcome to our newest members: members1004961949, W.T., Olanig, Ivanovic, akivela, Kolobunke, Gaznach1364, spiciz, Lexibont. New in last 24 hours: 72 Members, 1632 Posts". Below this is a blue banner that says "Welcome to the Forums at Psych Central." and provides instructions for new visitors. The main content area is titled "General" and contains a list of forum categories with their respective descriptions, sub-forums, and recent activity. The categories include: New Member Introductions, General Social Chat, Other Mental Health Discussion, Community Feedback & Technical Support, Kudos & Affirmations, General Q&A, Sanctuary for Spiritual Support, Relationships & Communication, Coping with Emotions, and The Sanity Score. Each category lists the last post, the number of threads, and the number of posts.

Forum	Last Post	Threads	Posts
New Member Introductions	Holla from Colombia by laduan Today 06:27 PM	35,266	241,926
General Social Chat	California? by albatud Today 06:49 PM	20,011	438,459
Other Mental Health Discussion	Are you asking for help when... by subia101a Today 06:43 PM	18,946	206,485
Community Feedback & Technical Support	New chat room software by buranzer Today 11:29 AM	5,514	30,878
Kudos & Affirmations	Hugs by tharaisa Today 12:05 PM	7,537	67,941
General Q&A	I am awfully lonely by patsifan Today 06:22 PM	1,947	14,299
Sanctuary for Spiritual Support	Please pray for by tharaisa Today 12:09 PM	2,845	31,891
Relationships & Communication	Marriage separation after... by cadroit Today 06:32 PM	20,128	185,580
Coping with Emotions	I'm done by albatud Today 09:55 PM	5,744	41,045
The Sanity Score	Just took the test by zeez Yesterday 02:22 PM	2,625	7,337

Figure 8.3. Psych Central Forums Home Page

8.4.2 Supporting active help providers to support others

As outlined in Chapter 1, there are numerous ways in which the coordinated delivery of mental health services requires improvement [40, 41]. There is not a single solution to the problems facing the mental health sector, but there are some areas such as e-mental health and peer support that have been identified as offering important potential to assist. The findings of this thesis suggest that, at the interface of these two areas, there are individuals who voluntarily provide peer support to thousands of people, and through which thousands more have the potential to benefit passively. There may be cost-effective and wide-reaching benefits to implementing policies that support these highly-engaged individuals in a manner that may enhance their impact.

In face-to-face services, the role of peer workers has been supported by the development of an accreditation scheme for peer work. For example, in Australia a nationally recognised qualification known as the 'Certificate IV in Mental Health Peer Work' has been funded and developed [42]. This accreditation scheme is designed for consumers and carers who are employed as peer workers. Given the potential influence of high-engaged users in MHISGs, consideration could be given to developing and offering a version of this course that is tailored to MHISG users who take on leadership roles as active help providers to other users within the ISG. Further research and design work would be required to inform the development, implementation and impact of such a course. Other approaches to supporting online peer leaders could also be investigated and debated. For example, pharmaceutical companies have recently begun sponsoring and publicising the proceedings of conventions for online health advocates [43].

8.5 Methodological strengths and limitations of the investigation

8.5.1 Strengths

The aim of this thesis was to develop a greater understanding of MHISG participation. Ultimately, there is little benefit to this knowledge unless it is linked to outcomes. However, the design of optimal interventions is facilitated by a scientific understanding of the nature of the processes that underpin them; this thesis has both advanced our understanding of the

nature of MHISGs and generated testable hypotheses for future research on both MHISG processes and outcomes.

The studies in this thesis have employed a range of methodologies, novel and conventional, and qualitative and quantitative. It has demonstrated the benefits of novel automated methods that can be employed on large-scale networks and which avoid the potential biases associated with selection effects in time-consuming manual analysis methods which necessarily rely on sampling a subset of data. Notably, the use of a topic-model to investigate the nature of content across the entirety of BlueBoard post content provided a comprehensive method of analysing the content that might be critical to the dynamics of the community. This allowed for detection of broad trends in the data including the novel finding that user conceptualisations of recovery may differ for high- and low-engaged users [4]. A follow-up study (in Chapter 7) using a qualitative manual content analysis also provided evidence of these differing conceptualisations of recovery. Thus, the use of differing methodologies to triangulate the findings served to increase the validity of the study conclusions. Similarly, converging evidence of awareness differences between high- and low-engaged users emerged in both the quantitative topic model analysis and the qualitative post analysis. This evidence of convergence in outcomes for the different methodologies suggests that topic modelling may be a valid exploratory technique for future research on MHISGs. Furthermore, the use of data collected from users during registration and the log data of their activity was also an important strength of the methods employed in this thesis. This data provided a more reliable perspective on community structure and user characteristics than previous depression ISG studies which to date have relied on survey data of a small subset of ISG users.

8.5.2 Limitations

This thesis involved a mixture of qualitative and quantitative methods in part to address the methodological limitations associated with each and to provide a means of validating the results by triangulating their respective findings. However, this overall approach and the conclusions that could be drawn from it, was limited by the restriction of the content analysis (Chapter 7) to a small sub-set of high-engaged users since there are so few of the latter.

Although this thesis employed systematic approaches and methods designed to overcome the potential for bias that was evident in previously published studies, it is possible that these methods have failed to elicit the underlying complexity of behaviour in MHISGs. Quantitative

studies, such as topic modelling and modularity analysis, are capable of detecting broad patterns but may not be sufficiently sensitive to reveal key details underpinning active participation. While, the qualitative analysis of initial posts did enable a more in-depth analysis, this analysis was confined to the user's initial post. Thus, although together the studies provided breadth and depth of knowledge, they did not necessarily provide a comprehensive understanding of an MHISG.

The possible analyses and hence conclusions in the studies in this thesis were limited by the available user registration data. The validity of data provided anonymously online cannot be guaranteed. Further, reliable data was not available for symptom severity, chronicity, diagnosed condition, attitudes to seeking professional help and stigma. Future research that incorporates these measures may provide a more nuanced or even an alternative perspective on the nature of participation in an MHISG.

Since the current studies were focused on a single MHISG, it is unclear if the findings from this thesis are applicable to other MHISGs and, moreover, online health communities broadly. Thus there is a need to replicate studies in the thesis in other MHISGs, particularly the study in Chapter 7. Again, the limitation regarding the quantity of available data for high-engaged users in their initial posts imposes restrictions on the degree to which the findings may be reliable. If the development of a tool for predicting user engagement is to be progressed, studies will require access to data from multiple MHISGs.

In Chapter 5, the modularity of the BlueBoard social network was inferred from co-participation in a thread and connections between users were weighted by the number of different threads in which both users participated. This methodology had some limitations. It may have underestimated the degree to which some high-engaged users communicated with each other in a single thread and overestimated the extent to which they connected with the rest of the community. Similarly, in Chapter 6, assumptions were made about the direction of communication in threads, with the user who initiated the thread being assumed to be the intended recipient of all communications (posts) made thereafter. Future analyses that are based on more refined models of directed communications may provide a more nuanced or alternative perspective on the nature of participation in an MHISG.

Finally, statistically, the power law distribution of engagement presented a challenge for the quantitative analyses throughout this thesis. Linear regression could not be used due to

violations of the statistical assumption of normality of residuals. Non-linear regression would have limited the interpretability of the output. However, the logistic regression and non-parametric tests that were undertaken instead reduced the power of the analyses and precluded a study of potential differences at the full range of the points in the spectrum of engagement.

8.6 Future research

8.6.1 Direct extensions

The current findings suggest a number of directions for future research with respect to (i) the nature and (ii) the outcomes of participation. Each is discussed in turn below.

(i) Nature of participation. The studies in this thesis require replication in other MHISGs in order to assess their reliability. In particular, there is a need to investigate if the differing conceptions of recovery between higher- and lower-engaged users are evident in other MHISGs, and if so to undertake follow-up studies to examine the nature of this difference. For example, studies should determine the extent to which the difference in conception of recovery is explicitly acknowledged by users in an MHISG, or whether the current findings are a function of some other important difference between users.

The empirical studies in this thesis focused on frequency of posts. For pragmatic reasons, a wide range of other participatory styles were not investigated. Developing a better understanding of other participation styles, including those identified by the systematic review in Chapter 2 may have promising applications (e.g., in the optimisation of the dissemination of health information). Further research is also required to investigate if there are other distinct participation styles not identified to date, and if so, to investigate their potential role in the delivery of tailored health promotion strategies.

(ii) Outcomes of participation. Studies are needed to examine a user's mental health and related outcomes as a function of their interactions within an MHISG. The Tripartite Model of MHISG Participation could provide a framework for the design of such investigations. It is neither feasible nor ecologically valid to randomly allocate users to groups in which they are

instructed to participate using different styles. However, it may be informative to track who is interacting with whom and correlate these occurrences with outcomes of participation, such as empowerment, hope, symptoms, stigma or attitudes towards seeking professional help. Such outcomes may be measured by very brief surveys following user interactions, with effects adjusted for initial status recorded at the time of registration. It would not be necessary to measure all of these outcomes in a single survey or a single person. Other important outcomes could be measured passively. For example the activity of users, such as click-throughs on links referring users to other professional sources of help, could be tracked online.

Furthermore, optimising MHISG service delivery will require systematic testing of the effects of various modifications to MHISGs to determine what works best and for whom. With respect to the modifications suggested in the Implications section (8.4) above (welcome messages and skills training for users), it should be noted that the desired mental health and behavioural outcomes may differ as a function of different sub-groups of users. For example, a successful manipulation of the welcome messages may increase the engagement of users with high awareness characteristics, whereas success in the case of low-awareness users may increase their professional help-seeking behaviour through referral information. Similarly, it might be hypothesised that a skills training intervention for high-engaged users will yield positive effects on personal recovery through the helper-therapy principle [44] with improvements to positive self-identity [45, 46], whereas low-engaged users who interact with the trained users, may experience greater improvements in hope, empowerment and willingness to seek help [47]. Finally, there is potential to test the effect of identifying and responding to ISG users employing automated methods such as those that detect symptoms from text [48] and provide automatic tailored feedback in response to users.

8.6.2 Broader implications for future research

As the granularity in the focus on the nature of participation in an online health community is refined, the conception of the community as an intervention in its own right diminishes and a conception of it as a lens through which to explore and influence health behaviour increases, paralleling our notion of face-to-face communities. However, from the perspective of a research enterprise, the online context presents a strong advantage for investigating health behaviour over the off-line context in that the data for all users is recorded by default. The ability to quantify and detect health behaviour and attitudes in this context thus provides a powerful tool to conduct research on topics that are difficult to study in offline behaviour. For

example, using Facebook social network data, a recent study has documented the spread of ‘emotional contagion’ through users of the website following the experimental manipulation of the proportion of positive and negative sentiment content in users’ newsfeeds [49]. Using the same principle, it may be possible to investigate the spread of other psychological phenomena and associated behaviours such as stigma and attitudes towards mental health help-seeking through social networks such as MHISGs. However, a major limitation of this approach is that the data, in its natural language format, must be processed before it can be used to infer these outcomes accurately. This barrier may be addressed in part by undertaking research which employs machine-learning techniques, such as those in the current thesis [4] to detect the phenomena of interest in the content of user interactions.

8.7 Conclusion

This thesis advances our understanding of MHISGs, illuminating the roles of users who may be luminaries for others, and progressing the model of user participation from the current simple active/passive dichotomy to a more sophisticated tripartite framework. The thesis provides evidence that posting frequency is not simply a reflection of a participant engaging in more or less of the same intervention, but rather is associated with fundamental differences in styles of participation. It was observed that higher posting frequency is generally associated with users engaging in mutual help with other higher-engaged users, and providing an active help provider role to lower-engaged active help seekers, even though higher- and lower-engaged users tended to differ in their approach to recovery. This new model and understanding provides a fertile ground for future research studies and methodologies, and may inform the work of policy makers and MHISG community managers tasked with optimising service delivery. There are exciting opportunities to further explore the potential of MHISGs.

8.8 References

1. Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM. From help-seekers to influential users: A systematic review of participation styles in online health communities. *Journal of Medical Internet Research*. 2015; 17(12): e271.
2. Griffiths KM, Carron-Arthur B, Reynolds J, Bennett K, Bennett A. User characteristics and usage of an open access moderated Internet support group for depression and other mental disorders: A prospective study. *Internet Interventions*. 2017; 7: 9-15.
3. Carron-Arthur B, Cooper M, Kazan D, Ali K, Griffiths KM. Are first impressions predictive of future engagement? Qualitative analysis and blinded judgements of users' initial post in a mental health Internet support group. In Preparation.
4. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Griffiths KM. What's all the talk about? Topic modelling in a mental health Internet support group. *BMC Psychiatry*. 2016; 16(1): 367.
5. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Cunningham JA, Griffiths KM. Community structure of a mental health Internet support group: Modularity in user thread participation. *JMIR Mental Health*. 2016; 3(2): e20.
6. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions*. 2014; 1(4): 165-8.
7. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. *Patient Education and Counseling*. 2008; 70(2): 193-8.
8. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. *Journal of Medical Internet Research*. 2010; 12(3): e34.
9. Lawlor A, Kirakowski J. Online support groups for mental health: A space for challenging self-stigma or a means of social avoidance? *Computers in Human Behavior*. 2014; 32: 152-61.
10. van Mierlo T, Hyatt D, Ching AT. Mapping power law distributions in digital health social networks: methods, interpretations, and practical implications. *Journal of Medical Internet Research*. 2015; 17(6): e160.
11. van Mierlo T, Li X, Hyatt D, Ching AT. Demographic and indication-specific characteristics have limited association with social network engagement: Evidence from 24,954 members of four health care support groups. *Journal of Medical Internet Research*. 2017; 19(2): e40.
12. Urbanoski K, Van Mierlo T, Cunningham J. Investigating patterns of participation in an online support group for problem drinking: a social network analysis. *International Journal of Behavioral Medicine*. 2016; Epub ahead of print.
13. Deetjen U, Powell JA. Informational and emotional elements in online support groups: a Bayesian approach to large-scale content analysis. *Journal of the American Medical Informatics Association*. 2016; 23(3): 509-13.
14. Wang X, Zhao K, Street N. Analyzing and predicting user participations in online health communities: A social support perspective. *Journal of Medical Internet Research*. 2017; 19(4): e130.
15. Wang X, Zuo Z, Zhao K. The evolution of user roles in online health communities: A social support perspective. Pacific Asia Conference on Information Systems; 2015.
16. Johnson SL, Safadi H, Faraj S. The emergence of online community leadership. *Information Systems Research*. 2015; 26(1): 165-87.
17. Hogg MA. A social identity theory of leadership. *Personality and Social Psychology Review*. 2001; 5(3): 184-200.

18. Dean J, Potts HWW, Barker C. Direction to an Internet support group compared with online expressive writing for people with depression and anxiety: a randomized trial. *JMIR Mental Health*. 2016; 3(2): e12.
19. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. *PLoS one*. 2012; 7(12): e53244.
20. Barak A, Grohol JM, Pector E. Methodology, validity, and applicability: A critique on Eysenbach et al. *British Medical Journal*. 2004; 328: 1166.
21. Ellis LA, Campbell AJ, Sethi S, O'Dea BM. Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. *Journal of CyberTherapy and Rehabilitation*. 2011; 4(4): 461-7.
22. Tomasino KN, Lattie EG, Ho J, Palac HL, Kaiser SM, Mohr DC. Harnessing peer support in an online intervention for older adults with depression. *The American Journal of Geriatric Psychiatry*. 2017; Epub ahead of print.
23. Barak A, Dolev-Cohen M. Does activity level in online support groups for distressed adolescents determine emotional relief. *Counselling and Psychotherapy Research*. 2006; 6(3): 186-90.
24. Nimrod G. Online depression communities: Does gender matter? *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*. 2012; 6(1): Article 6.
25. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. *American Journal of Psychiatry*. 2002; 159(12): 2062-8.
26. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences*. 2016; 25(02): 113-22.
27. Wright, K. B. Communication in health-related online social support groups/communities: A review of research on predictors of participation, applications of social support theory, and health outcomes. *Review of Communication Research*. 2016; 4: 65-87.
28. Nimrod G. The membership life cycle in online support groups. *International Journal of Communication*. 2012; 6: 1245-61.
29. Borgatti SP, Everett MG. Models of core/periphery structures. *Social Networks*. 2000; 21(4): 375-95.
30. Collier B, Kraut R. Leading the collective: social capital and the development of leaders in core-periphery organizations. *arXiv*. 2012: 1204.3682.
31. Nielsen J. Participation inequality: lurkers vs contributors in Internet communities 2014 [cited 2014 05/06/2014]. Archived at <http://www.webcitation.org/6Q7EwEncA>. Available from: <http://www.nngroup.com/articles/participation-inequality/>.
32. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. *Journal of Medical Internet Research*. 2011; 13(4): e122.
33. van Uden-Kraan C, Drossaert C, Taal E, Seydel E, van de Laar M. Self-reported differences in empowerment between lurkers and posters in online patient support groups. *Journal of Medical Internet Research*. 2008; 10(2): e18.
34. Davidson L, Chinman M, Kloos B, Weingarten R, Stayner D, Tebes JK. Peer support among individuals with severe mental illness: A review of the evidence. *Clinical Psychology: Science and Practice*. 1999; 6(2): 165-87.
35. Grohol JM. Psych Central Forums 2010 [cited 2017 15/06/2017]. Archived at <http://www.webcitation.org/6rE3X1aF0>. Available from: <https://forums.psychcentral.com/new-member-introductions/163968-active-participation-increases-value-your-time-here.html>.

36. Griffiths KM, Reynolds J, Vassallo S. An online, moderated peer-to-peer support bulletin board for depression: User-perceived advantages and disadvantages. *JMIR Mental Health*. 2015; 2(2): e14.
37. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. *BMC Psychiatry*. 2011; 11(1): 88.
38. Australian Government Department of Health. A national framework for recovery-oriented mental health services: Guide for practitioners and providers. Canberra: Australian Government Department of Health, 2013.
39. Young C. Community management that works: how to build and sustain a thriving online health community. *Journal of Medical Internet Research*. 2013; 15(6): e119.
40. National Mental Health Commission. Contributing Lives, Thriving Communities - Review of Mental Health Programmes and Services. Canberra: National Mental Health Commission, 2014.
41. World Health Organization. Comprehensive Mental Health Action Plan 2013–2020. Geneva: World Health Organization; 2013.
42. Australian Government. CHC42912 - Certificate IV in Mental Health Peer Work (Release 2) 2017 [cited 2017 15/06/2017]. Archived at <http://www.webcitation.org/6rEC3OBFP>. Available from: <http://training.gov.au/training/details/CHC42912>.
43. Janssen Global Services. HealthEVoices Conference 2017 [cited 2017 15/06/2017]. Archived at <http://www.webcitation.org/6rEBfqGR> . Available from: <https://www.healthvoices.com/>.
44. Riessman F. The "helper" therapy principle. *Social Work*. 1965; 10(2): 27-32.
45. Andresen R, Caputi P, Oades L. Stages of recovery instrument: development of a measure of recovery from serious mental illness. *Australian & New Zealand Journal of Psychiatry*. 2006; 40(11-12): 972-80.
46. Davidson L, Strauss JS. Sense of self in recovery from severe mental illness. *Psychology and Psychotherapy: Theory, Research and Practice*. 1992; 65(2): 131-45.
47. Corrigan PW. Impact of consumer-operated services on empowerment and recovery of people with psychiatric disabilities. *Psychiatric Services*. 2006; 57(10): 1493-6.
48. Karmen C, Hsiung RC, Wetter T. Screening Internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. *Computer Methods and Programs in Biomedicine*. 2015; 120(1): 27-36.
49. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*. 2014; 111(24): 8788-90.