

Predicting User Response and Support Activities in Virtual Health Support Communities

Joseph A. Manga
University of Texas RGV
joseph.manga01@utrgv.edu

Bin Wang
University of Texas RGV
bin.wang@utrgv.edu

Abstract

Despite growing emphasis on the factors affecting different types of supports users receive from virtual health support communities (VHSC), theoretical knowledge on how social awareness capabilities determine the extent of the support received is yet to be investigated. Adopting social awareness theory and using data collected from the COVID-19 support community on a large VHSC platform, we apply linguistic analysis to measure the impacts of three social awareness variables — social sensitivity, social insight, and social communication — on users' response and support behaviors. The ordinary least square regression results show that social insight significantly influenced the number of replies to a post. In addition, results from the negative binomial regression also indicate that social sensitivity and social communication significantly predicted the number of support votes and thanks votes a user's post received. The findings reveal some important research and practical implications on the need to facilitate social awareness in VHSC forums.

1. Introduction

With the increase in social media attention and interaction, the focus on self has blurred individuals' ability to consider the social needs and situations of others especially in virtual health support communities (VHSCs). Although recent studies have emphasized the need to reexamine individuals' voluntary supply of support and help in VHSCs using social capital [10] and structural capital [3], the degree and extent to which individuals in these VHSCs are willing to help is yet to be properly understood and investigated. There is a growing need to examine whether patients actually receive adequate supports and responses in terms of the quantity of information/replies and votes needed to address their health, informational, and emotional concerns. This current study addresses this concern from the social awareness theory perspective, which unveils a unique opportunity in investigating people's

ability to understand others relative to referent self; thus, considering the provisions of help to be more unidirectional rather than reciprocal. The blend between individuals' social capabilities (which are social awareness skills that help people understand others, appreciate their views, and contribute to a community or the society) with the new tools like virtual communities that encourage interactivity are opening a whole new era of using social media not only to benefit oneself but more so, to place greater importance on the needs of others. Consequently, while extant research has informed our understanding of the determinants of the types of support that are provided in virtual health communities, there is an increasing challenge to predict how much support people actually receive to their communication exchanges in these virtual health communities. This is important because people place value not only in the content of the help but also in the amount of help they receive. When the help and support an individual receives are sufficient, it builds satisfaction and continued usage [15] of the services on the social media platform. However, if the help and support to patients' worries are inadequate to address such needs, they begin to lurk [21]. Therefore, the fundamental question addressed in this study is:

- *Do individuals' social awareness capabilities predict user response behaviors in virtual health support community forums?*

Specifically, this current study aims to investigate how individuals develop capabilities through social awareness to improve their understanding of their peers in virtual health support communities, which prompt their selfless acts of providing support and replies needed by their counterparts. To better predict the amount of support to individuals' online posts, this study leverages social awareness theory [7], which comprises of three dimensions (social sensitivity, social insight, and social communication). With data sourced from the recently created Coronavirus community on a large VHSC platform, we performed an ordinary least

squares regression and a zero-inflated negative binomial regression model to estimate the hypotheses. Findings reveal that with the exception of social insight, individuals' social sensitivity and social communication expressed in users' posts significantly predicted the number of replies to a user's post and the total number of support and thanks votes a post receives. Findings suggest that individuals' ability to comprehend and communicate effectively with others in social settings can increase activeness in the community, thereby, improving the support and help the patients need to cope with their health challenges.

This study makes some important contributions to theory and practice. First, our findings showed that social awareness theory can be extended to specific online settings to understand how individuals react to others' health challenges. Specifically, individuals can express their social awareness capabilities via their posts to help patients cope and deal with the recent COVID-19 disease. With the social stigma associated with COVID-19, it is imperative for people to develop social skills needed to reduce the impact of this disease. The findings in this current study suggest that patients can feel safe and exonerated from shaming—for a pandemic that is posing threats to everyone—when individuals develop insights, sensitivity, and good communication skills towards their peers facing this disease. Second, this study provides evidence that supporting others deal with their health conditions requires one to show sympathy and empathy by identifying with others' feelings even when you need help yourself. Third, people receive more support when users' social communication skills are in play. Mastering the act of generating effective responses to influence other's behaviors contributes significantly to increase the extent of support patients receive.

Studying about the extent to which people react to other's posts and provide supports in VHSCs is relevant because it shows some signals of activeness on the platforms; and the platform acts as a medium of high internet traffic, which can benefit advertising firms who could use it as an opportunity to market their products in order to generate more revenues. OHSC platform managers can identify users' social awareness skills through the tone of their words expressed in their posts and responses. Management should encourage users who sympathized with and adequately respond to others' needs and concerns and discourage users who make fun of or a joke of other's challenges.

2. Background review

2.1. Virtual health support communities

Virtual communities in general are noted for the increased information support, emotional support, and companionship activities that take place within these forums [10] [3]. Additionally, knowledge and information sharing are also key activities that characterize these online platforms [27] [3]. There are growing numbers of studies in this area that have shown that for users in virtual communities to receive the support they need, they have to share some personal information (e.g., [26]). Since the activities and the users of the online communities are different, subsets of the general online communities are beginning to gain steam and drawing research attention. In particular, virtual health support communities (VHSC) focuses on health-related issues where users with different health conditions join to share and seek for support [12] [18] [28]. "*A health-related virtual community is a website that enables people to create profile pages, post information, and participate in online discussions of health-related information,*" ([12], p. 55). Prior literature has revealed that structural, relational, and cognitive social capitals emotional support and companionship activities between patients in VHSCs [10], structural social capital can indirectly influence participants' health literacy and attitudes in VHSC [3], and VHSCs can encourage and improve community participation and response behaviors [11] [29] [19]. Furthermore, information from VHSCs facilitate patients' decision to switch between online and offline medical services [14].

Considering the unique characteristics that VHSCs provide and the vast contributions that research on VHSCs have made to identify and explain the antecedents of the types of support that patients receive, a significant concern that still needs to be addressed is to provide a theoretical explanation of what drives patients' enthusiasms to post and to respond to a threaded post and how much support do patients actually receive? This study provides answers as it seeks to predict the volume of supports and responses that users in these VHSCs are willing and able to generate in the course of their discussions. This is important because the vibrancy of the interaction between users in VHSCs directs internet traffic to the platform/website, which can be useful medium for targeted marketing by platform management and third-party firms. Moreover, the huge amounts of data generated on these platforms can help medical professionals to research certain types of diseases and treatment options. To address our concern, knowledge from social awareness theory is leveraged to understand the determinants that make users of VHSCs to be full of energy and enthusiasm by posting, replying, and providing support to other users' online discussion posts.

2.2. Social awareness theory

Social awareness refers to “an individual’s ability to understand people, social events, and the processes involved in regulating social events. The emphasis on interpersonal understanding as the core operation in social awareness indicates that this construct is a cognitive component of human competence” ([7], p. 18). Social awareness is a multidimensional construct that includes social sensitivity, social insight, and social communication. Social sensitivity is the ability to correctly describe the meaning of a social entity or incident, that is, to place oneself in other people’s positions to know how others feel [7]. Social insight refers to the ability to know how to intervene effectively and influence the behaviors of others [7]. Social communication is the ability to interact aptly with others, including problem-solving communications [7]. Social awareness is one component of a larger model of personal competence that also includes emotional competence, physical competence, conceptual intelligence, and practical intelligence. Social awareness has been used to study team performance activities and results showed that being aware of one’s social group increases members’ satisfaction [1]. Furthermore, social awareness manifests as empathy: an emotional competence that considers other’s needs and concerns. Empathy is a social sensitivity attribute that is vital in improving relationships and in shaping response behaviors [5].

3. Model and hypotheses development

As shown in Figure 1, we hypothesize that the three dimensions of social awareness are critical in predicting the total number of replies and votes a user receives from the interactions and communication exchanges in virtual health support communities.

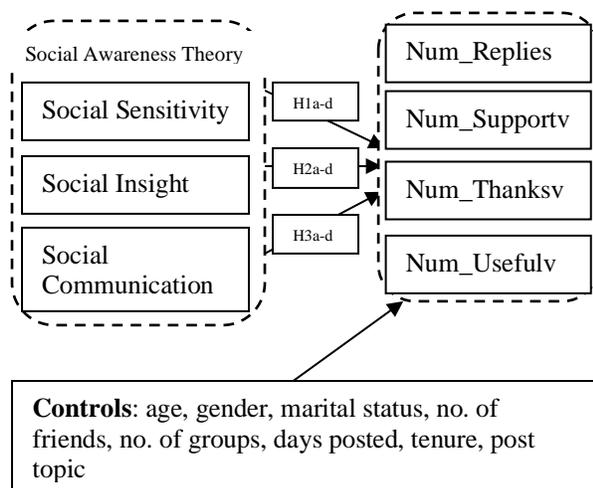


Figure 1: Research Model

3.1. Social sensitivity

Following the prior literature, we define *social sensitivity as an individual’s ability to identify oneself with the feelings of others in a VHSC discussion forum*. In a social setting, it is an individual’s ability to read situations accurately and comprehend other people’s feelings and viewpoints [7]. To effectively demonstrate social sensitivity towards an event, one needs keen attention and rightly interpret the social event [16]. In VHSC forums, patients come with different issues and health conditions. When a user initiates a post, others reply to the post and provide votes of support, thanks, and useful. However, the mechanism behind the respondents’ responses are driven by social sensitivity of the author’s original post. A post with higher social sensitivity will resonate more with the readers and get more responses and votes. This means that the discloser feels some empathy for and identifies with the needs of other patients on the platform; hence, readers are more likely to provide responses to the initial post. However, if one’s social sensitivity is low, which means the author is not considerate of the needs of others, the probability of receiving replies and support to the post will be low. Hence,

Hypothesis 1a-d: Social sensitivity is positively associated with the (a) number of replies, (b) number of support votes, (c) number of thanks votes, and (d) number of useful votes provided to users’ posts in virtual health support communities.

3.2. Social insight

Prior study defines social insight as the ability to know the processes underlying social events and to make evaluative decisions about such events [7]. *Social insight in our context refers to an individual’s ability to accurately understand the needs of others and provide suitable responses to others in a VHSC discussion forum*. This ability is demonstrated by an individual’s level of understanding of different social classes, friendships, expectations, and values of others, and of how to respond in social gatherings [19]. It involves decision skills of selecting and making an appropriate response to others [16]. In VHSC platforms, individuals join with varied needs and expectations, which must be met for them to be satisfied. Thus, providing the necessary support whether informational, relational, or emotional [10] that patients need requires a high degree of insights in understanding what they are going through

and choosing an appropriate response to meet the patients' expectations. People who demonstrate high social insights through their posts will more accurately diagnose other people's need for sociability, or friendship, support and will more likely receive responses and help from their peers. Thus, the volume of responses will be significant when the social insight is high as evidenced in the original threaded post. So,

Hypothesis 2a-d: Social insight is positively associated with the (a) number of replies, (b) number of support votes, (c) number of thanks votes, and (d) number of useful votes provided to users' posts in virtual health support communities.

3.3. Social communication

Social communication has been defined as an individual's ability to generate effective responses to influence other's behaviors [7]. It involves the capacity to generate effective strategies of dealing with others, to accurately convey one's feelings to others, and an understanding to solve problems and influence others. *In this study, social communication refers to an individual's ability to effectively convey their thoughts with others in a virtual health support community forum.* Individuals with good social communication skills tend to be more influential and they intervene in most group discussions. In VHSC forums, a person with high communication ability could be identified as power users or thread initiators; and study has shown that the number of responses posted by a power user is significantly high [13]. Consequently, posts by users with better social communication skills will generate more replies or receive more votes of support, thanks, and useful. Consequently,

Hypothesis 3a-d: Social communication is positively associated with the (a) number of replies, (b) number of support votes, (c) number of thanks votes, and (d) number of useful votes provided to users' posts in virtual health support communities.

4. Methodology

4.1. Setting and data source

Data was sourced from a large VHSC platform specifically, the Coronavirus support community at <https://www.inspire.com/groups/covid-19/> between March and April 2020. The Coronavirus support group was chosen because of its sensitivity to public scrutiny and public shaming. As a result, there is a high need for social support for users facing stigmatized health

conditions in order to reduce the effects of this stigma on patients. The virtual health support platform in general hosts a large number of support groups or communities involving users' interactions via posting and responding to each other on varied health topics such as, "Symptoms", "Testing", "Risk factors", "Treatment", "Prevention", etc. The key dependent variables were directly measured from features on the platform.

4.2. Variables and measures

4.2.1. Dependent variables. These include the natural log of the number of replies $\ln(\text{Num_Replies})$, the number of support votes (Num_Supportv), the number of thanks votes (Num_Thanksv), and the number of useful votes (Num_Usefulv) a user post receives from other users. These variables were directly observed from the VHSC platform. To reduce skewness and heteroscedasticity in the number of replies, we took the natural log transformation. The other dependent variables' values are small, and the raw counts were used.

4.2.2. Independent variables. All the independent variables were measured using sentiment scores from the linguistic inquiry word count (LIWC) text analytics [22]. The measurements of the focal independent variables are as follows. *Social sensitivity* measured by obtaining the *Analytic* thinking score – logical and more formal way of thinking or personal or informal thinking; some stem words associated with this variable are responsiveness, reactivity, thoughtfulness, understanding, empathy, feeling. *Social insight* measured by obtaining *Insight* score – speaking with high expertise and confidence or more humble and tentatively; some stems words include discernment, judgement, vision, sharpness, acuity, savvy. *Social communication* measured by obtaining the *Tone* score – which represents positive and upbeat speaking or sadness and hostility in speaking; some stem words include conveying, strength, firmness, devoted, determined. Prior research have made use of the LIWC dictionary lexicon-based approach to extract and measure key variables in their studies (see [8]). Table 1 presents how the variables were operationalized.

4.2.3. Control variables. This study controlled for *gender*, *age*, *marital status*, user activity in terms of *number of friends* (this controls for the number of friends a user has as the higher the number of friends, the more the replies and votes): this was normalized by taking the natural log transformation; *days since user posted* (this controls for the number of days the post has

been posted or been on the platform since older posts tend to have more replies and votes, calculated as date of data collection minus date the message was posted); *number of communities/groups* users belong to (this controls for the number of communities a user belongs to since posting in many groups attract more friends who tend to provide more replies and votes to their posts); *topic of post* (we generated and used eight dummy variables *DummyPTOi* for the nine different topics that

users posted on, where $i = 2, \dots, 8$); *longevity/experience* [25], which is the period of time the user has been on the platform were used as controls in the empirical analysis (the longer the user has been on the platform, the more experienced they are in understanding others and hence, attracting more replies and votes; calculated as the natural log of postdate minus date user joined the platform).

Table 1: Variables and operationalization

Variable Description		Operationalization
Dependent variables		
Replies	ln(Num_Replies) (lnNRE)	The natural log of the number of replies a user post receives.
Support votes	Num_Supportv (NSV)	The number of “support” votes a user post receives.
Thanks votes	Num_Thanksv (NTV)	The number of “thanks” votes a user post receives.
Useful votes	Num_Usefulv (NUV)	The number of “useful” votes a user post receives.
Independent Variables		
Social sensitivity	Social_Sens (SSE)	Analytic thinking score obtained from sentiment analysis.
Social insight	Social_Insi (SIN)	Insight score obtained from sentiment analysis.
Social communication	Social_Comm (SCO)	Tone score obtained from sentiment analysis.
Control variables		
Longevity on the site	ln(Experience) (lnEXP)	The natural log of the number of days a user has been a member in the group.
Topic of post	DummyPTOi (PTO)	Vector of eight dummy variables for nine heading or topic categories including “general,” “isolation/quarantine,” “news and updates,” “prevention,” “research,” “risk factors,” “testing,” “symptoms,” and “treatment.”
Days posted	DaysPosted (DPD)	The number of days the message was posted on the forum.
User’s friends	ln(No.Friends) (lnNFR)	The natural log of the number of friends the author of the post has.
User’s groups	No.Groups (NGP)	The number of groups the author of the post belongs to.
User’s age known	DummyAge (DAG)	1 if the user’s age available; 0 otherwise.
Actual age	AAGE	The post author’s actual age if available; 0 otherwise.
User’s gender known	DummyGender (DGE)	1 if the post author’s gender was known; 0 otherwise.
Actual gender	AGEN	1 if the post author’s gender was male; 0 if female or unknown.
User’s marital status	MStatus (MST)	1 if the user was married; 0 otherwise.

4.2.3. Sample. The sample includes data collected on users’ textual posts, number of replies for each post, and numbers of support, thanks, and useful votes for the different posts. We cleaned the data to remove duplicate observations and possible outliers (posts by users the researchers identified as group leaders). The data set

comprises of 177 unique user posts collected for different users who participated in discussions on different topics concerning the recent COVID-19 pandemic in the VHSC forum. The user names were randomly assigned unique user ids for privacy reasons.

Table 2 presents the descriptive statistics of the variables used in the analysis.

Table 3 presents the correlations of the main variables in the model. The correlation values did not show high pairwise correlations. Hair et al. [9] suggest variance inflation factors (VIF) less than 10 are indicative of inconsequential collinearity. The highest VIF was 6.18, below the acceptable level of 10 and a mean VIF of 3.06. Thus, the correlations do not pose collinearity problems.

Table 2: Descriptive statistics (N=177)

	Mean	S.D.	Min	Max
SSE	66.05	26.68	1.00	99.00
SIN	2.46	2.23	0.00	10.45
SCO	39.78	33.12	1.00	99.00
NSV	2.41	4.39	0.00	37.00
NTV	2.75	5.35	0.00	35.00
NUV	1.62	3.28	0.00	23.00
lnNRE	1.52	1.13	0.00	7.06
AAGE	28.62	30.02	0.00	84.00
DAG	0.51	0.50	0.00	1.00
AGEN	0.11	0.32	0.00	1.00
DGE	0.60	0.49	0.00	1.00
lnEXP	6.46	2.19	0.00	8.39
lnNFR	1.85	2.04	0.00	7.15
NGP	2.44	1.98	0.00	5.00
MST	0.21	0.41	0.00	1.00
DPD	13.42	7.51	1.00	40.00

Table 3: Correlations

Variable	1	2	3	4	5	6	7
1. SSE	1						
2. SIN	-0.11	1					
3. SCO	0.09	0.23**	1				
4. NSV	-0.12	0.10	0.17*	1			
5. NTV	0.14	-0.0002	0.25**	0.79***	1		
6. NUV	0.13	-0.003	0.19*	0.80***	0.90***	1	
7. lnNRE	0.08	0.055	0.09	0.49***	0.43***	0.484**	1

Variable Inflation Factor (VIF) and 1/VIF

	SIN	SCO	NSV	NTV	NUV	lnNRE
1.22	1.10	1.18	4.00	6.07	6.18	1.67
0.82	0.91	0.85	0.25	0.17	0.16	0.60

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; N = 177

5. Empirical analysis and results

In the analysis, we first leverage text analytics (specifically, sentiment analysis using LIWC) to obtain and measure the key social awareness variables. Because we transformed NRE by taking the natural log of the number of replies, it is no longer a count variable

but continuous. Thus, the ordinary least squares (OLS) regression technique was used to model the number of replies $\ln(\text{Num_Replies})$ as shown in Equation 1 below. For the rest of the count outcome variable (votes), the generalized linear models (GLM) with negative binomial regression (NBREG) technique [6] was employed since our dependent variables are count of the various votes (support, thanks, and useful) given to a user's post. The selection of this technique is best suited for modeling count outcome variables; usually for count outcome variables that are overdispersed, that is, when the conditional variance exceeds the conditional mean [4]. Overdispersion in the data was confirmed using the alpha and the likelihood ratio tests [2]. The NBREG was preferred over the zero-inflated negative binomial (ZINB) [23] since the zero inflation was found to be insignificant. The model used to test the hypotheses related to the number of replies is:

$$\ln(\text{Num_Replies}_i) = \beta_0 + \beta_1 \text{Social_Sens}_i + \beta_2 \text{Social_Insi}_i + \beta_3 \text{Social_Comm}_i + \beta_4 \ln(\text{Experience}_i) + \beta_5 \ln(\text{No.Friends}_i) + \beta_6 \text{No.Groups}_i + \beta_7 \text{DaysPosted}_i + \beta_8 \text{DummyAge}_i + \beta_9 \text{AAge}_i + \beta_{10} \text{DummyGender}_i + \beta_{11} \text{AGen}_i + \beta_{12} \text{MStatus}_i + \text{DummyPTO}_i \beta_{13} + \varepsilon_i \quad (1)$$

The models used to test the hypotheses related to the votes are:

$$P(Y_{ij}=y) = NB(\mu_{ij}), \quad (2)$$

where $j=(1, 2, 3)$ represents the three support, thanks, and useful votes, Y_{ij} is the votes for the i th post, $NB(\cdot)$ is a negative binomial distribution with mean μ_j and variance $\mu_j(1 + \alpha \mu_j)$ and

$$\mu_{ij} = \exp(\gamma_{0j} + \gamma_{1j} \text{Social_Sens}_{ij} + \gamma_{2j} \text{Social_Insi}_{ij} + \gamma_{3j} \text{Social_Comm}_{ij} + \gamma_{4j} \ln(\text{Experience}_{ij}) + \gamma_{5j} \ln(\text{No.Friends}_{ij}) + \gamma_{6j} \text{No.Groups}_{ij} + \gamma_{7j} \text{DaysPosted}_{ij} + \gamma_{8j} \text{DummyAge}_{ij} + \gamma_{9j} \text{AAge}_{ij} + \gamma_{10j} \text{DummyGender}_{ij} + \gamma_{11j} \text{AGender}_{ij} + \gamma_{12j} \text{MStatus}_{ij} + \text{DummyPTO}_{ij} \gamma_{13j} + \varepsilon_{ij}). \quad (3)$$

Two models were performed to obtain the results for each outcome variable as shown in Table 4. The first columns for each outcome variable include only the control variables and the second columns includes both the effects of the main independent variables and the controls. The dummies of the post topics (PTO) vector variable were also included in the analysis.

Using the model in Equation 1, first, we present the results of the ordinary least squares (OLS) regression used in modeling $\ln(\text{Num_Replies})$ or $\ln\text{NRE}$ in Table 4. The OLS results show that social insight was negative and significant in predicting the natural log of the number of replies to a user's post, contrary to our

hypothesis ($\beta = -0.007$, $p < 0.1$). About 24% of the variance in the natural log of the number of replies was explained by the predictors, an improvement from

21.6% with just the control variables. Effects of social sensitivity and social insight were not significant.

Table 4: Hypothesis testing results (N=177).

Variables	lnNRE		NSV		NTV		NUV	
Constant	2.380*** (0.440)	2.647*** (0.496)	1.177** (0.581)	1.889*** (0.664)	2.087*** (0.577)	1.536** (0.663)	0.652 (0.591)	0.795 (0.693)
SSE		-0.003 (0.004)		- 0.016*** (0.005)		0.005 (0.004)		-0.002 (0.005)
SIN		-0.007* (0.004)		0.008 (0.049)		-0.082 (0.053)		-0.062 (0.053)
SCO		-0.001 (0.003)		0.003 (0.003)		0.012*** (0.004)		0.004 (0.004)
AAGE	0.017** (0.008)	0.017** (0.008)	-0.008 (0.011)	-9.87e-05 (0.011)	-0.007 (0.012)	-0.006 (0.012)	-0.008 (0.012)	-0.008 (0.012)
DAG	-1.258** (0.509)	-1.313** (0.518)	0.433 (0.670)	-0.235 (0.670)	-0.061 (0.723)	-0.102 (0.714)	-0.213 (0.720)	-0.255 (0.739)
AGEN	0.421 (0.278)	0.437 (0.282)	-0.095 (0.349)	0.081 (0.341)	0.519 (0.340)	0.555 (0.339)	-0.221 (0.358)	-0.176 (0.362)
DGE	-0.285 (0.178)	-0.298 (0.181)	-0.179 (0.242)	-0.263 (0.240)	-0.368 (0.238)	-0.406* (0.233)	-0.069 (0.245)	-0.080 (0.246)
lnEXP	-0.002 (0.044)	0.013 (0.047)	0.024 (0.058)	0.073 (0.059)	-0.003 (0.055)	-0.022 (0.055)	0.052 (0.062)	0.058 (0.064)
lnNFR	0.046 (0.049)	0.042 (0.050)	0.162*** (0.062)	0.178*** (0.061)	0.175*** (0.059)	0.203*** (0.059)	0.255*** (0.060)	0.263*** (0.062)
NGP	0.017 (0.059)	0.009 (0.061)	0.052 (0.071)	0.011 (0.071)	-0.155** (0.072)	-0.157** (0.071)	-0.109 (0.076)	-0.110 (0.076)
MST	-0.365 (0.255)	-0.324 (0.270)	-0.633* (0.336)	-0.590* (0.340)	0.133 (0.331)	-0.222 (0.337)	-0.318 (0.346)	-0.456 (0.368)
DPD	0.024** (0.011)	0.024** (0.011)	0.019 (0.013)	0.023* (0.012)	0.039*** (0.012)	0.036*** (0.012)	0.051*** (0.013)	0.049*** (0.013)
PTODummies	Included	Included	Included	Included	Included	Included	Included	Included
R ²	0.216	0.238						
Adj. R ²	0.132	0.141						
Prob > F	0.001	0.001						
lnalpha			0.153 (0.179)	0.033 (0.185)	0.062 (0.176)	-0.049 (0.182)	-0.234 (0.246)	-0.244 (0.247)
alpha			1.166	1.033	1.064	0.952	0.791	0.783
L.L.			-329.38	-323.05	-328.11	-322.23	-253.67	-252.65
Wald Chi ²			50.90	63.56	80.19	91.96	83.43	85.48
Prob>Chi ²			0.001	0.001	0.001	0.001	0.001	0.001
Pseudo R ²			0.072	0.090	0.109	0.125	0.141	0.145

Notes: N = 177. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Second, using the models represented by Equations 2 and 3 for the rest of the count outcome variable NSV, NTV and NUV, the NBREG models were performed and the overall results indicate a good fit for the number of support votes (NSV), thanks votes (NTV) and useful votes (NUV) with highly significant log likelihood ratios ($p < 0.01$ for Wald Chi²) as shown in Table 4. Social sensitivity (SSE) negatively predicted the

number of support votes - NTV ($\beta = -0.016$, $p < 0.01$), while social communication (SCO) positively predicted the number of thanks votes - NTV ($\beta = 0.012$, $p < 0.01$) to a post. Surprisingly, none of the main independent variables was found to be significant in predicting the number of useful votes (NUV) a post receives.

The results of this study also reveal that some control variables were important factors in the prediction

model. Specifically, users who disclosed their age (DAG), the specific topic a user is posting on (PTO), and the number of days since a user posted (DPD) were significant factors in predicting the number of replies (lnNRE) to a post. Meanwhile, the number of friends a user has (lnNFR), the number of days since a user posted (DPD), and the specific topic a user is posting on (PTO) were the three main controls that were important in the prediction model for the number of support, thanks, and useful votes (NSV, NTV, NUV) a user post receives. Although one would expect that the greater the number of friends a user on the platform has, the higher the number of replies to their post. However, this was not the case as the results did not reveal any significant findings.

6. Discussion and contribution

It was expected that social awareness dimensions will act as influential predictors of active user interaction in virtual health support communities. Specifically, a user's social insight significantly predicted the volume of replies that patients get. However, a negative relationship was found. This suggests that users' posts tend to receive less replies when their messages reveal high accuracy of understanding the needs of their peers. In other words, if the posts have low social insight, users may still receive many replies and votes to their posts. This is the same with social sensitivity, which also negatively predicted the number of support votes contrary to what was hypothesized. These are important social skills because individuals with high social sensitivity and social insight can make excellent social relationships among group members and increase the collective group cognitive intelligence to perform a variety of tasks [17].

One possible explanation for the unexpected results is that the negative relationships could be due to the fact that users who visit the platform, come to primarily seek for emotional and informational supports themselves. It is likely that their posts portray low sensitivity and insights towards others due to the anxiety and distress they brought with on the platform. We also examined the data to see if there were similarities in the posts that received the greater number of replies and votes. We found that there seems to be some pattern with the users' posts that bagged in more replies and the votes came from users who expressed uncertainty and anxiety about the COVID-19 pandemic. As such, users are unlikely to express high social sensitivity and social insights in their initial (very first) posts. It is probable that such feelings and understanding could be conveyed in their subsequent posts or replies to other's posts; hence,

suggesting that further analysis on user's subsequent posts is needed to validate this claim. One last reason for these negative relationships is that it is possible that the measures used are too general to assess social sensitivity and social insight, thus, requiring more specific measures for the variables.

On a more favorable note, this study found that social communication positively predicted the number of thanks votes as expected. This suggests that other users appreciate more when initiators of posts demonstrate positive emotions and more upbeat emotional tone in their posts. Since users who visit the online health support community come with challenging and threatening health conditions, they will tend to appreciate posts that address their emotional needs to help them change their health behaviors. Prior studies supplement the findings of this study. In fact, besides the word-of-mouth type of communication, which has been shown to improve interaction in online communities [2], other study has also revealed that users' participation and contributions to the success of online communities are greatly influenced by new types of communication features best described as "ballot box communication" such as tagging and rating other users' posts, which reflect the opinions of individuals [26].

In general, users with high senses of social capabilities will generally understand others due to their increased awareness of others' feelings, appreciate others' views, provide meaningful contributions to benefit the communities or groups they belong to, and increase participation in the community [20]. Additionally, individuals with high senses of social capabilities will tend to display more benevolent behaviors by responding and replying more rather than reciprocating. This is because they see it as good citizenship behaviors and they also find it more fulfilling and pleasing. Furthermore, there will be more user engagements and community participation in online discussions when individuals show more responsible social awareness behaviors [24]. Users' attitudes could change when they notice their peers are more engaging and responsible in considering other people's needs and views above theirs. In the section that follows, some implications to research and practice are discussed with future research directions.

6.1. Implications for research and practice

This study has implications both to theory and practice. For research, first, through social awareness, we can understand why users show high engagements and participation in virtual health support community platforms. Second, we applied social awareness theory to the virtual health support community platforms and specifically to the recent COVID-19 support group and

showed that the blaming and shaming of patients with such disease can be reduced through apt demonstrations of social awareness capabilities in social settings. Third, we used predictive modeling to investigate the adequacy of the support that people receive from their online activity. Studies have informed our understanding of the types of support but not how much support is received. Thus, this current study pushed that limit to throw more evidence on how much support patients need to alter their health behaviors. Fourth, based on our predictive model, we showed that the amount of support that individuals receive is as important as the quality of the support. By applying predictive modelling, this study makes a method contribution to research.

To practice, first, managers should design rewards systems (in the form of badges and points) to measure users' social capabilities through the way they observe, listen, and provide appropriate responses to others on the platform. Second, from the results, it is clear that users with low social sensitivity and social insight still receive high number of replies and votes to their posts. Online health support community platform managers can potentially boost users' sensitivity and insight skills by assuring and making them understand that there are similar individuals in the group who face the same health conditions as theirs and who need more support and help as they do. This will cause users to demonstrate more social sensitivity and insight in their messages, showing that one is not just focusing on self, but is aware and respect other peers on the platform. Third, platform managers should understand that insensitive users could ruin community interconnectedness and participation among users. So, a suggested way of enhancing users' social awareness skills is by introducing social awareness games as part of the platform design to improve participants sensitivity and insight skills. For example, users can be asked to identify different facial expressions about what they think the person illustrated in pictures feel or think. Fourth, given that social communication is an important component in the model, the platform design could incorporate machine learning tools that provide suggests to users on the tone of their messages. Lastly, platform managers could design quizzes to test and identify individuals' social awareness skills based on the stem words used in measuring social sensitivity, social insight, and social communication.

6.2. Limitations and future research

This study has some limitations and provides opportunities for further research. First, a limited data set did not provide a big picture of the effects of the social awareness dimensions on the number of replies and votes a user post gets. A larger sample is needed to validate the results in future studies. Second, the results

need to be interpreted with caution as the setting may not be generalizable. Future research could explore different settings other than online health support communities to understand the implications of the proposed model. Third, the measurements of the theoretical variables were mainly from the sentiment scores using linguistic inquiry word count (LIWC). Although this current study found some significant results with these measurements, the results can be improved if alternative measurements are used. Thus, future research could examine the proposed model from a survey-based study and use other settings. Lastly, this current study used a cross-sectional data set to examine the user response activities and controlled for the effect of time (user's longevity on the platform). However, is possible that users may become less anxious and more certain over a long period, therefore, expressing more social awareness skills towards others in their posts. So, further research should perform a longitudinal study to examine if the effects four social awareness theory factors could change over time.

6.3. Conclusion

This current study drew from social awareness theory to extend the literature discussion on user activities in virtual health support communities. With the exception of the number useful votes (NTV) that was not significantly predicted by any of the predictor variables, findings reveal that insight predicted number of replies (lnNRE), social sensitivity predicted number of support votes (NSV), and social communication predicted number of thanks votes (NTV). These factors are important capabilities for online users to have in order to increase the level response and support activities that take place in the virtual health platforms. These capabilities are necessary and important qualities that should be demonstrated by platform users who understand the feelings and needs of others. Since most people who come on the platform to seek for some form of support, it is expedient for them to receive sufficient support not only in quality but also in quantity. The results also confirm and provide evidence to support prior findings in literature on online health support communities, which serve as means for activeness in user participation and continual usage of the platform services by the members. Lastly, in order for the social capabilities to be sustained, users need to be aware of themselves, which provides an opportunity for future research on how to consistently manage and maintain self-awareness towards self and social awareness towards others in online settings.

7. References

- [1] Bender, L., G. Walia, K. Kambhampaty, K.E. Nygard, and T.E. Nygard, "Social sensitivity correlations with the effectiveness of team process performance: An empirical study", *Proceedings of the ninth annual international conference on International computing education research*, (2012), 39–46.
- [2] Cameron, A.C., and P.K. Trivedi, *Regression analysis of count data*, Cambridge university press, 2013.
- [3] Chen, L., A. Baird, and D. Straub, "Fostering participant health knowledge and attitudes: An econometric study of a chronic disease-focused online health community", *Journal of Management Information Systems* 36(1), 2019, pp. 194–229.
- [4] Faraj, S., S. Kudaravalli, and M. Wasko, "Leading collaboration in online communities.", *Mis Quarterly* 39(2), 2015.
- [5] Goleman, D., "An EI-based theory of performance", *The emotionally intelligent workplace: How to select for, measure, and improve emotional intelligence in individuals, groups, and organizations 1*, 2001, pp. 27–44.
- [6] Gomes, M.J.T.L., F. Cunto, and A.R. da Silva, "Geographically weighted negative binomial regression applied to zonal level safety performance models", *Accident Analysis & Prevention* 106, 2017, pp. 254–261.
- [7] Greenspan, S., "Defining childhood social competence: A proposed working model.", *Advances in special education*, 1981.
- [8] Gunarathne, P., H. Rui, and A. Seidmann, "When social media delivers customer service: Differential customer treatment in the airline industry", *MIS Quarterly* 42(2), 2018, pp. 489–520.
- [9] Hair, J.F., R.E. Anderson, R.L. Tatham, and W.C. Black, "Multivariate Data Analysis", 3rd ed, *Macmillan Publishing Company, New York*, 1995.
- [10] Huang, K.-Y., I. Chengalur-Smith, and A. Pinsonneault, "Sharing Is Caring: Social Support Provision and Companionship Activities in Healthcare Virtual Support Communities", *MIS Quarterly* 43(2), 2019, pp. 395–423.
- [11] Kordzadeh, N., and J. Warren, "Communicating Personal Health Information in Virtual Health Communities: A Theoretical Framework", *2014 47th Hawaii International Conference on System Sciences*, IEEE (2014), 636–645.
- [12] Kordzadeh, N., and J. Warren, "Communicating personal health information in virtual health communities: An integration of privacy calculus model and affective commitment", *Journal of the Association for Information Systems* 18(1), 2017, pp. 1.
- [13] Kordzadeh, N., C. Zhechao Liu, Y.A. Au, and J.G. Clark, "A multilevel investigation of participation within virtual health communities", *Communications of the Association for Information Systems* 34(1), 2014, pp. 26.
- [14] Li, H., Y. Fang, K.H. Lim, and Y. Wang, "Platform-Based Function Repertoire, Reputation, and Sales Performance of E-Marketplace Sellers", *MIS Quarterly* 43(1), 2019, pp. 207–236.
- [15] Liao, R., R. Kishore, and M. Lee, "Users' Continued Usage of Online Healthcare Virtual Communities: An Empirical Investigation in the Context of HIV Support Communities", 2019.
- [16] McFall, R.M., "A review and reformulation of the concept of social skills.", *Behavioral assessment*, 1982.
- [17] Meslec, N., I. Aggarwal, and P.L. Curseu, "The insensitive ruins it all: Compositional and compilational influences of social sensitivity on collective intelligence in groups", *Frontiers in psychology* 7, 2016, pp. 676.
- [18] Mpinganjira, M., "Precursors of trust in virtual health communities: A hierarchical investigation", *Information & Management* 55(6), 2018, pp. 686–694.
- [19] Mpinganjira, M., "Willingness to reciprocate in virtual health communities: the role of social capital, gratitude and indebtedness", *Service Business* 13(2), 2019, pp. 269–287.
- [20] Ning Shen, K., and M. Khalifa, "Exploring multidimensional conceptualization of social presence in the context of online communities", *Intl. Journal of human-computer interaction* 24(7), 2008, pp. 722–748.
- [21] Nonnecke, B., J. Preece, and D. Andrews, "What Lurkers and Posters Think of Each Other.", *HICSS*, Citeseer (2004).
- [22] Pennebaker, J.W., R.L. Boyd, K. Jordan, and K. Blackburn, *The development and psychometric properties of LIWC2015*, 2015.
- [23] Simons, J.S., D.J. Neal, and R.M. Gaher, "Risk for marijuana-related problems among college students: An application of zero-inflated negative binomial regression", *The American Journal of Drug and Alcohol Abuse* 32(1), 2006, pp. 41–53.
- [24] Wang, Y.-C., R. Kraut, and J.M. Levine, "To stay or leave? The relationship of emotional and informational support to commitment in online health support groups", *Proceedings of the ACM 2012 conference on computer supported cooperative work*, (2012), 833–842.
- [25] Wasko, M.M., and S. Faraj, "Why should I share? Examining social capital and knowledge contribution in electronic networks of practice", *MIS quarterly*, 2005, pp. 35–57.
- [26] Yang, L., and B.C. Tan, "Self-disclosure on online social networks: motives, context feature, and media capabilities", 2012.
- [27] Yang, X., G. Li, and S.S. Huang, "Perceived online community support, member relations, and commitment: Differences between posters and lurkers", *Information & management* 54(2), 2017, pp. 154–165.
- [28] Zhang, X., S. Liu, X. Chen, L. Wang, B. Gao, and Q. Zhu, "Health information privacy concerns, antecedents, and information disclosure intention in online health communities", *Information & Management* 55(4), 2018, pp. 482–493.
- [29] Zhou, J., "Factors influencing people's personal information disclosure behaviors in online health communities: A pilot study", *Asia Pacific Journal of Public Health* 30(3), 2018, pp. 286–295.