Helping Each Other Quit Online: Understanding User Engagement and Real-life Outcomes of the r/StopSmoking Digital Smoking Cessation Community

ALESSIO DE SANTO, HEG Arc, HES-SO, University of Applied Sciences Western Switzerland, Switzerland
ARIELLE MORO, BRUNO KOCHER, and ADRIAN HOLZER, University of Neuchâtel, Switzerland

Despite decades of prevention, tobacco addiction is still a widespread health concern responsible for around 8 million deaths per year. Existing digital smoking cessation solutions such as social media are becoming increasingly popular and represent a novel approach to find community support. However, little is known about how they affect smoking behavior. This research aims to understand what motivates people to join online communities and how their participation affects their attitudes and behaviors. To do so, this article conducts an in-depth analysis of the popular Reddit r/StopSmoking thread through three complementary studies. Using the transtheoretical model and the uses and gratification theory, Study 1 aims at understanding the link among motivation factors, engagement, and outcomes through a user survey. Study 2 aims at understanding the engagement by analyzing the content of 10 years of user interaction data. Study 3 attempts to gain further knowledge of interactions by examining the reaction of the community to a crisis situation such as that of the recent COVID-19 pandemic. Findings convey the fact that participation in such communities has a favorable impact on the change process toward quitting. Results show that providing social support to others is the biggest contributing factor for participating in the community. User interactions analysis confirmed that survey responses were accurate reflections of actual user activity. Regarding the impact of the COVID-19 crisis, results suggest that it increased levels of stress and depression in the community while decreasing active engagement, indicating that there may be opportunities for improvement in dealing with tough situations.

CCS Concepts: • **Applied computing** → **Consumer health**; *Health care information systems*; • **Information systems** → *Social networking sites*; • **Social and professional topics** → Health information exchanges;

Additional Key Words and Phrases: Smoking cessation, digital support, data analysis, reddit

ACM Reference format:
Alessio De Santo, Arielle Moro, Bruno Kocher, and Adrian Holzer. 2023. Helping Each Other Quit Online: Understanding User Engagement and Real-life Outcomes of the r/StopSmoking Digital Smoking Cessation Community. *ACM Trans. Soc. Comput.* 6, 1, Article 1 (April 2023), 30 pages.
https://doi.org/10.1145/3564745

Authors' addresses: A. De Santo, HEG Arc, HES-SO, University of Applied Sciences Western Switzerland, Espace de l'Europe 21, Neuchâtel, Neuchâtel, Switzerland, 2000; email: alessio.desanto@unine.ch; A. Moro, B. Kocher, and A. Holzer, University of Neuchâtel, A.L.-Breguet 2, Neuchâtel, Neuchâtel, Switzerland, 2000; emails: {arielle.moro, bruno.kocher, adrian.holzer}@unine.ch.

2469-7818/2023/04-ART1 $15.00
https://doi.org/10.1145/3564745

1 INTRODUCTION

According to the World Health Organization, tobacco use is the starting point for many non-communicable diseases and is responsible for over 8 million deaths annually [64]. Most smokers, who are aware of the dangers of tobacco, would like to quit but need help to do so [63]. Simple behavioral change interventions can considerably reduce the premature deaths of tobacco users [63]. Past research has investigated which interventions can effectively reduce the premature deaths of tobacco users. According to primary care guidelines, brief advice, behavioral support, pharmacotherapy, and abstinence evaluation are the only four interventions providing strong evidence of efficacy [58]. Unfortunately, these evidence-based interventions are still largely underused [54]. For instance, face-to-face counseling, which is the most effective way to help smokers to quit [22], has only low participation rates [38] and is also not affordable globally [61]. Reaching smokers with this efficient intervention will require novel approaches to enhance the effectiveness of existing cessation interventions, and to increase their adoption.

To support individuals in adopting healthier behaviors, digital artifacts are increasingly available [13, 31, 51]. Social networks [6], social support [62], and social integration [9] appear to play important roles in smoking behavior and can potentially provide helpful complementary support for smoking cessation [22]. Among available digital resources, social media are becoming increasingly popular for people in search of health information and support [11, 50]. On these platforms, the digital community, i.e., the group of people interacting on the platform, can potentially provide 24/7 behavioral support and brief advice through messaging. However, evidence on the specific factors that enable digital communities to support tobacco cessation are still limited [8, 57]. This leads to the overarching research question of this article:

RQ: How do digital smoking session communities support users in their quitting process?

To answer this research question, this article provides an in-depth analysis of an online smoking cessation community through three distinct yet complementary studies. The smoking cessation community chosen for this purpose was the r/StopSmoking thread on Reddit, which has over 100k users, 115k posts and 750k comments, all publicly available.

Study 1, aims at understanding the link among motivation factors, engagement, and outcomes through a user survey. This study contributes to the literature by outlining a novel research model, which is based on the **transtheoretical model (TTM)** [45] as well as the uses and gratification theory [4, 29]. Our model aims to predict online engagement and eventually the progress in smokers' behavioral process of change toward becoming an ex-smoker. We further expand the socialization factor of the uses and gratification theory by making the distinction between providing social support and seeking social support. We validate this model using survey data from 169 users of r/StopSmoking.

Study 2 aims at understanding the engagement on r/StopSmoking to discover the main type of information available, the proportion of social support provided, the evolution of active users, the evolution of active engagement and the smoking status of the contributing authors. Such a data-driven approach allows us to corroborate the data reported by users in Study 1. To do so, we extracted the interactions over 10 years and manually tagged 5,000 random r/StopSmoking interactions to test and train different multi-label algorithms enabling the classification of 10 years of r/StopSmoking interactions.

Study 3 aims at understanding r/StopSmoking reaction to a special crisis situation such as that of the COVID-19 pandemic. To do so, Study 3 takes a multimodal analytics approach combining survey results assessing psychometric variables and activity traces of r/StopSmoking interactions during the period 2017 to 2020. The particular psychometric variables measured include general health aspects (e.g., **Patient Health Questionnaire (PHQ)** [32]), as well as stress (e.g., **Perceived Stress Scale (PSS)** [12, 59]) and anxiety aspects (e.g., Spielberger **State-Trait Anxiety Inventory**

(STAI) [35, 55]). To understand the impact of the COVID-19 crisis on interaction, a seasonalized time-series regression [23] is used to predict the levels of interactions one would have observed without the crisis.

2 LITERATURE REVIEW

This section reviews the literature relevant to all three studies of our work. Any specific literature for individual studies is covered in the appropriate study sections.

2.1 Evidence-based Interventions

Four interventions for smoking cessation have been found to be supported by evidence [58]: (1) pharmacotherapy, (2) brief advice, (3) behavioral support, and (4) abstinence evaluation. Examples of pharmacotherapy include nicotine replacement therapy and the use of bupropion or varenicline to assist patients with nicotine withdrawal. A brief advice is 5–10 minutes of advice to encourage smokers to improve their health by quitting their smoking habit, primarily by triggering a cessation attempt. Behavioral support includes: self-help material, peer group meetings and health professional counseling. *Self-help information* can support patients without outside help. When self-help is personalized, it is even more effective [49]. With *peer group meetings*, smokers who attempt to quit meet regularly and provide each other with support and encouragement. *Health professional counseling* generally consists of one-on-one face-to-face appointments between a medical professional and a smoker. Enhancing the motivation to stop smoking through behavioral support has been identified as an important aspect of the overall treatment for tobacco addiction [49]. Finally, abstinence evaluation is the confirmation of abstinence through either self-reporting or objective measures such as biochemical markers or clinical tests.

Digital artefacts can potentially be designed to support several evidence-based interventions. For instance, digital communities have been found to contribute to significant long-term positive health outcomes through their information-providing role (self-help material) [15, 37]. However, it is not yet clear how engagement in such communities is linked to the actual process of behavior change. Digital interventions also make it easier for peers to connect and support one another. Unfortunately, research results as to their effectiveness are not yet conclusive [56]. A better differentiation of social support concepts and causal pathways requires further investigation to be able to demonstrate the effective value of social relationships in increasing smokers' likelihood of cessation [62].

2.2 Digital Communities for Smoking Cessation

Several studies have analyzed online smoking cessation communities [18, 34, 57, 66]. Some studies investigated the effectiveness of health behavior interventions on online social communities, finding none, very modest or ambivalent evidence of efficacy [34, 66]. Others used observational rather than interventional approaches to understand peer-generated content and interactions [7, 18]. From this perspective, emerging results suggest that peer support is helpful in avoiding smokers relapsing [57] or to motivate them to quit [58]. Researchers argue that digital communities can be perceived as a "safe space" for smokers to talk about day-to-day challenges, cravings, or relapses [18]. Furthermore, the relative anonymity of the Internet can facilitate discussion and mutual support [18]. A wide variety of peers can be available at any time to provide help and support through various activities, such as sharing information, sympathizing, cheering, coaching, or celebrating. Some of these peers, potentially further along the line, can be considered as *expert patients*. They can provide firsthand experience about how to cope, what to expect, and how things feel [66].

Digital interventions can be seen as potentially supporting several evidence-based interventions. For instance, digital communities have been found to contribute to significant long-term positive

Table 1. Transtheoretical Model Stages of Change

Stages of Change	***Description***
Precontemplation (Stage 1)	No intention to take action within the next 6 months
Contemplation (Stage 2)	Intends to take action within the next 6 months
Preparation (Stage 3)	Intends to take action within the next 30 days and has taken some behavioral steps in this direction
Action (Stage 4)	Changed overt behavior for more than 6 months
Maintenance (Stage 5)	Changed overt behavior for more than 6 months
Termination (Stage 6)	No temptation to relapse and 100% confidence

health outcomes through their information-providing role (*self-help material*) [15, 37]. However, it is not yet clear how engagement in such communities is linked to the actual process of behavior change. Digital interventions can also be seen as potentially supportive *peer group meetings*. Unfortunately, research results are not yet conclusive [56]. A better differentiation of social support concepts and causal pathways is awaiting further investigations to demonstrate the effective value of social relationships in improving smokers likelihood of cessation [62].

3 STUDY 1: UNDERSTANDING HOW A DIGITAL COMMUNITY SUPPORTS SMOKING CESSATION

The goal of the first study is to understand the link between motivational factors, community engagement, and behavioral outcome, (i.e., quitting smoking). Quitting smoking seems to be more a process than a single act following a decision as illustrated by the director of the Cipret Neuchâtel, a regional anti-tobacco agency, who explains that "on average it takes around seven attempts before someone manages to quit for good."

A useful theoretical model, matching this idea of process, is the TTM [45]. This model, based on stages, differs from many other behavioral theories that are based on so-called continuum models. According to continuum models, interventions could be applied in any order, or even simultaneously, and they do not include any notion of progression [53]. Stage models imply that different interventions are appropriate at different stages of health behavior change [53], making for instance TTM a frequently used model for smoking cessation intervention [46]. Stage models have been criticized [24, 60], arguing that the notion of stages might be flawed or circular, in that the stages are not genuinely qualitative, but they still embed this notion of progression, which goes beyond the sole variation of intention to include the action and post-action spectrum.

TTM hypothesizes that this change occurs in six distinct steps, also called *stages of change* (Table 1).

The TTM further suggests a set of 10 *processes* mediating the progress between stages (see Table 2). Empirical integration [47] suggests that, in the early stages, smokers rely on cognitive, affective, and evaluative processes to progress through the stages. In later stages, smokers work more on commitments, conditioning, contingencies, environmental controls, and support for progressing toward maintenance or termination.

We argue that one of the important factors in assessing the impact of online communities is participant (i.e., user) engagement. User engagement in a digital community can be defined as the different activities that users perform in an online community. These activities can be divided into active contributing activities (e.g., posting messages, reacting to a comment, sharing a video) and passive consuming activities (e.g., viewing content, visiting a page) [19, 36]. Based on these definitions and the aforementioned literature, we make the following hypotheses:

- **H1:** Overall, engagement in digital smoking cessation communities is positively linked to the process of change.

Table 2. Processes of Change That Mediate Progression between the Stages of Change

Process of Change	***Description***
Consciousness raising *(Stage 1 → 2)*	Increasing awareness via information, education and personal feedback about the healthy behavior
Dramatic relief *(Stage 1 → 2)*	Feeling fear, anxiety or worry because of the unhealthy behavior, or feeling inspiration and hope when they hear about how people are able to change to healthy behaviors
Environmental reevaluation *(Stage 1 → 2)*	Realizing the negative impact of the unhealthy behavior or the positive impact of the healthy behavior on one's proximal social and/or physical environment
Self-reevaluation *(Stage 2 → 3)*	Realizing that the behavior change is an important part of one's identity as a person
Self-liberation *(Stage 4 → 5)*	Making a firm commitment to change
Helping relationships *(Stage 5 → 6)*	Seeking and using social support for the healthy behavior change
Counterconditioning *(Stage 5 → 6)*	Substitution of healthier alternative behavior and cognition for the unhealthy behavior
Reinforcement management *(Stage 5 → 6)*	Increasing the rewards for the positive behavior change and decreasing the rewards of the unhealthy behavior
Stimulus control *(Stage 5 → 6)*	Removing reminders or cues to engage in the unhealthy behavior and adding cues or reminders to engage in the healthy behavior
Social liberation *(no specific Stage)*	Realizing that the social norms are changing in the direction of supporting the healthy behavior change

- **H1a:** Active engagement in digital smoking cessation communities is positively linked to the process of change.
- **H1b:** Passive engagement in digital smoking cessation communities is positively linked to the process of change.

To understand the motivational factors influencing engagement in digital communities, the *uses and gratifications theory* is widely relied upon [29], even though it was originally developed to examine how and why individuals use and adopt mass media in their daily lives [4]. As depicted in Table 3, this theory describes four motivational factors to predict engagement in digital communities [29, 36, 42]: *information-seeking*, *status-seeking*, *entertainment*, and *socialization*.

While the uses and gratifications theory already stands as a valid means of examining the motivations sought and obtained by users of online communities [29, 36, 42], recent literature is still calling for contributions to further expand this approach by, for instance, examining new needs and motivations [36]. To expand this model, one possibility is to integrate recent findings about behavior change [36]; in this case, the evidence is that giving support is even more useful than receiving support [16, 17]. Since smoking cessation is a process in which one is not immune from relapsing, providing support can have a twofold positive impact: for people receiving the support but also for people providing it. To reflect this dual aspect of seeking and providing support, we would suggest expanding the socialization factor into two relevant subfactors: (1) *Providing social support* and (2) *Seeking social support*. Providing social support is measuring users' motivation to support others in their smoking cessation process, which would represent "expert" users who are further ahead on the withdrawal journey and are willing to provide firsthand experience and

Table 3. Motivational Factors to Participate in Online Smoking Cessation Communities

Motivational factor	*Description*
Information seeking	Seeking and obtaining useful information is one of the primary motivations for Internet use [39]. Information seeking has been positively associated with social networks participation in past studies [29, 48]. Online communities represent a wealth of information, providing smokers with self-help material.
Status-seeking	Improving one's social status has been shown to be a strong motivating factor from a uses and gratifications perspective in studies of social network usage [29, 36].
Entertainment	The need for pleasurable, emotional and aesthetic experiences has been found to be a strong motivation factor in Internet and social network usage [36, 42]. Studies have indicated that reading and sharing content may also meet entertainment needs for users [14].
Socialization	Previous studies have found socialization as strong motivations among users [36, 42, 48]. Building and maintaining social contacts are among the most prevalent reasons for participation in online communities [36, 42]. Smokers may also use such channels to share about their quitting journey and to exchange peer support.

support [66]. Seeking social support aims to measure a user's motivation to engage in online communities to exchange with others to receive support. These observations lead to the following hypotheses:

- **H2:** Uses and gratification motivational factors are positively linked to engagement in online smoking cessation communities.
- **H3:** Seeking and providing social support are motivational factors that increase participation in online smoking cessation communities.

This study focuses on a particular digital smoking cessation community, the popular Reddit social media platform, and in particular its r/StopSmoking subreddit, which is one of the largest and most active communities dedicated to smoking cessation (more than 100k users, 115k posts, and 750k comments). A subreddit can be seen as a shared forum where users can post messages, reply with comments, and vote messages up or down. Each Reddit user, also known as a redditor, can join r/StopSmoking and ask for or give advice, share stories, or encourage someone who is trying to quit. The following sections provide a description of the survey, a presentation of the model and method used to analyze the data, and a description of the participants, through key data.

3.1 Method

3.1.1 Survey. This study employed a survey designed to assess individuals' motivation to participate in r/StopSmoking, their level of engagement, and the perceived influence of the community on their process of behavior change. The survey was validated by the University of Neuchâtel ethics committee and asked for informed consent of participants. Respondents were anonymous and could stop the study at any time. Besides asking for informed consent, demographics, and some descriptive facts in the first section, the survey aimed at measuring engagement, motivation, and process of change.

- *Measuring engagement*: The second section focused on engagement in r/StopSmoking. Engagement was measured by asking participants how frequently they performed the following types of activities: visiting, reading, posting, commenting, and voting. The first two activities are considered passive engagement, whereas the last three are instances of active engagement. These activities were measured using a five-point scale: "never," "almost never," "occasionally/sometimes," "almost every time," and "every time." Frequency of visiting was measured through a question asking for the average monthly number of visits with possible answers ranging from "less than 1" to "30 or more."

- *Measuring motivation*: The third section of the survey measured the motivation to participate in r/StopSmoking. Motivation to participate was measured through four factors, according to the uses and gratification approach: information-seeking, status-seeking, socialization, and entertainment. The information-seeking, status-seeking, and entertainment factors were directly taken and adapted from prior instruments employed in uses and gratifications research [29, 36, 42]. The socialization factor was also adapted from previous research, but the questions were linked to the novel factors *providing social support* or *seeking social support* according to the direction of the interaction. For instance, one of the questions used for the *providing social support* factor asked whether the respondent participated in the r/StopSmoking subreddit to help others. In contrast, one of the questions used for *seeking social support* factor asked if the respondent participated in the r/StopSmoking subreddit to gain peer support from others. All factors were measured with multiple questions, using five-point Likert scales.
- *Measuring process of change*: The fourth section of the survey focused on the behavior change process. The process of change was measured through the individual's accomplishment of the various processes mediating the progression between the stages of change of the TTM process of change (see Table 2). For instance, to measure whether r/StopSmoking helped people to move from precontemplation to contemplation, we inquired if r/StopSmoking allowed users to accomplish the corresponding processes of change: consciousness raising, dramatic relief, and environmental reevaluation. For instance, consciousness raising accomplishment was measured by asking participants if the digital community allowed them to increase their awareness via information, education, and personal feedback about smoking cessation (see Table 2). For each process, the accomplishment was measured with the same question structure, and individuals could then answer through a five-point Likert scale. Each stage of change was finally measured as a formative construct—based on the various processes that would allow them to progress beyond it—giving indicators measured through five-point Likert scales. As there are no clear processes allowing people to move on from the preparation and termination stages, they were not included in our model.
- *Participants*: Users of the r/StopSmoking subreddit were invited to participate in the survey through links posted directly to the site. The survey was completed on a voluntary basis and no compensation was given to respondents. A total of 173 responses were collected from February 11 to April 11, 2020. To maintain the visibility of the invitation among other posts, invitations to participate in the survey were randomly re-posted 18 times throughout the data collection period. After preliminary analyses and data preparation to remove incomplete surveys and surveys with inconsistencies in the control questions, 169 participants were included in the analysis (85 males, 83 females, and 1 preferred not to say; mean age 34). The average age when participants started to smoke was 16.6 years old. The average number of years of smoking was 16.1 (min 1 – max 45). Respondents smoked on average 16.8 cigarettes per day before quitting.

3.1.2 Model and Data Analysis. **Partial least squares (PLS)**, a variance-based **structural equation modeling (SEM)** analysis technique, was used for assessing our model and our

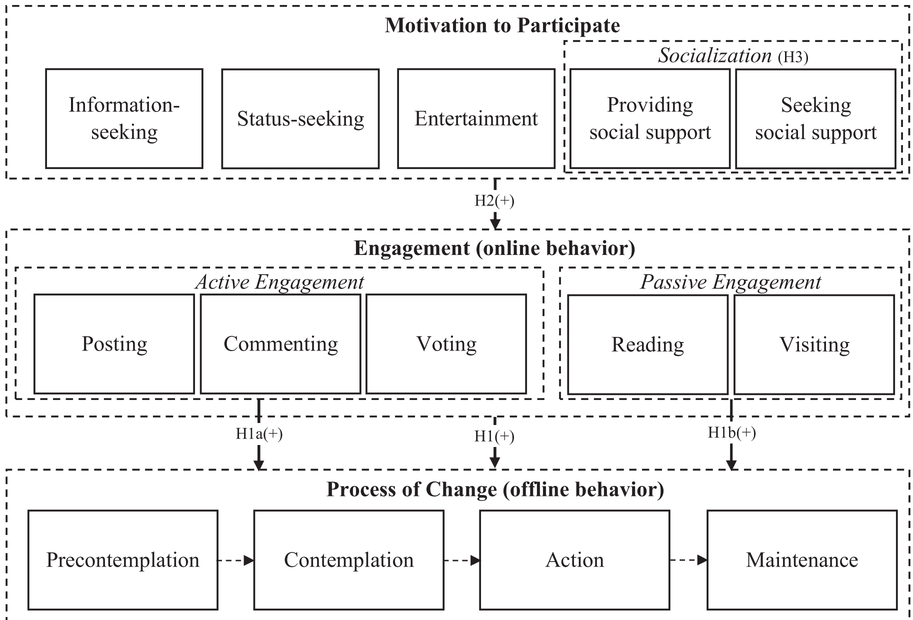


Fig. 1. Research model.

hypotheses. PLS is an increasingly popular technique in IS research to analyze explanation and prediction of IS phenomena [20, 52, 65]. Central to PLS is the path model, a diagram that displays the hypotheses and variable relationships to be estimated in an SEM analysis [5].

The construction of our path model (Figure 1) includes three main dimensions: (1) *Motivation to Participate*, referring to motivational factors influencing the participation in the digital community; (2) *Engagement (online behavior)*, referring to engagement in the digital community; and (3) *Process of Change (offline behavior)*, referring to the process of behavior change. To evaluate our model and test our hypotheses we used a three-stage approach. First, as advocated by Hair et al. [20], we evaluated the reliability, validity, and significance of our model. Second, we analyzed the overall results, focusing on a macro view of the model to test hypotheses H1 and H2. Finally, we tested hypotheses H1a, H1b, and H3 with an in-depth path analysis from a micro view of the model. SmartPLS was used as the analysis tool.

To evaluate the reliability of our *reflective constructs* (i.e., information-seeking, status-seeking, entertainment, providing social support and seeking help), we used **composite reliability (CR)** and **average variance extracted (AVE)** as indicators. As shown in Table 4, the CR of all the constructs was greater than 0.7 and the AVE greater than 0.5; hence, these constructs are considered reliable [20].

To evaluate the convergent validity of the reflective construct, we considered outer loadings and the AVE of the indicators [21]. The outer loadings of all our reflective variables were above 0.7 (the standard threshold [21]), apart from one indicator in status-seeking and another in entertainment that were above 0.6. As this study is an exploratory research, we decided to keep indicators between 0.4 and 0.7, as recommended by Hair et al. [21]. To measure the discriminant validity of our reflective constructs, we measured the Heterotrait-Monotrait Ratio [20]. The rule of thumb accepts values lower than 0.85 for conceptually distinct constructs and below 0.90 for conceptually similar constructs. As shown in Table 5, all values were lower than 0.85, demonstrating the discriminant validity of our constructs.

Table 4. Evaluation of Reflective Constructs

	rho_A	*CR*	*AVE*	*Cronbach's Alpha*
Information-seeking	0.860	0.890	0.730	0.822
Status-seeking	0.583	0.765	0.527	0.553
Entertainment	0.711	0.793	0.568	0.646
Prov. social support	0.880	0.940	0.887	0.873
Seek. social support	0.712	0.874	0.776	0.711

Table 5. Heterotrait-Monotrait Ratio

	Entertainment	*Information-seek.*	*Prov. social support*	*Seek. social support*	*Status-seek.*
Entertainment					
Information-seeking	0.384				
Prov. social support	0.323	0.152			
Seek. social support	0.585	0.527	0.755		
Status-seeking	0.414	0.298	0.686	0.611	

To validate our *formative constructs*, we measured the **variance inflation factor (VIF)**, defined as the reciprocal of the tolerance. The VIFs of our formative constructs indicating values lower than 5 (Table 6) exclude potential collinearity problems [21]. Weights express a formative indicator's relative importance in forming the construct. Significance indicates whether formative indicators truly contribute to forming the construct. The results presented in Table 6 depict the indicator's weight and significance for each formative construct. All the indicators of the motivation to participate construct were highly significant ($p < 0.001$). Concerning the engagement indicators, the frequency of posting and voting were not significant ($p > 0.05$) in the determination of individuals' active engagement. In the same way, consciousness raising was not found to be a significant determinant in the users' progression from precontemplation to contemplation stage. As recommended by Hair et al. [21], we verified that, for outer loadings of non-significant indicators found in our study, all values were high (>0.5) and indicators were eventually retained, bearing in mind that such indicators had an absolute and not relative importance. To assess the high-level hypotheses (H1 and H2), we extracted latent variable scores of motivation to participate, engagement, and process of change constructs.

3.2 Results

For each hypothesis investigated in Study 1, the sections below present the specific analysis and results obtained. The outcome of the analysis is presented in Figure 2.

3.2.1 Engagement in Digital Smoking Cessation Communities Is Positively Linked to the Process of Change (H1, H1a, H1b). Figure 2 shows a significant and positive link between engagement and process of change. The more users are engaged in the community, the higher their process of change. H1 is supported.

Table 6. Evaluation of Formative Constructs

Construct	*Indicator*	*Outer Weight*	*t-value*	*p-value*	*VIF*
Motivation to Participate	Information-seeking	0.259	4.812	<0.001	1.284
	Status-seeking	0.251	7.090	<0.001	1.377
	Prov. social support	0.340	9.452	<0.001	1.783
	Seek. social support	0.328	14.609	<0.001	2.129
	Entertainment	0.245	8.598	<0.001	1.343
Active Engagement	Posting	0.303	1.238	0.216	1.871
	Commenting	0.584	2.561	0.010	1.908
	Voting	0.324	1.745	0.081	1.219
Passive Engagement	Reading	0.613	2.799	0.005	1.145
	Visiting	0.602	2.845	0.004	1.145
Precontemplation	Consciousness raising	0.198	1.133	0.257	1.343
	Dramatic relief	0.415	2.993	0.003	1.236
	Environm. reevaluation	0.666	5.247	<0.001	1.234
Maintenance	Counterconditioning	0.302	3.115	0.002	1.426
	Helping relationships	0.476	4.880	<0.001	1.180
	Reinforcement managem.	0.374	3.907	<0.001	1.257
	Stimulus control	0.253	2.321	0.020	1.303

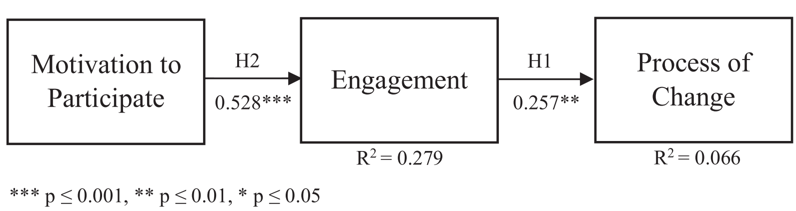


Fig. 2. Overall results from macroscopic view of the model.

Engagement explains 6.6% of the total variance of the overall process of change. The quite poor influence of r/StopSmoking overall engagement on the process of change is not surprising, as it is difficult to exert an influence on every stage of the process at the same time. It is for this reason that we tested hypotheses H1a and H1b at a lower level of granularity and analyzed the influence of both active and passive engagement on each stage independently. Our results, shown in Figure 3, show that both passive and active engagement have a significant positive influence on the process of change. Thus, both H1a and H1b are supported. It should be noted that active and passive engagement influence the stages of change differently. Active engagement influenced only processes of change moving on maintenance in a strongly significant manner, while passive engagement influenced processes of change moving on precontemplation and maintenance with high levels of significance (see Figure 3). However, engagement did not significantly influence processes of change moving from contemplation to action, neither for moving from action to maintenance.

3.2.2 Uses and Gratification Motivational Factors Are Positively Linked to Engagement in Online Smoking Cessation Communities (H2). Figure 2 shows that motivation to participate influences

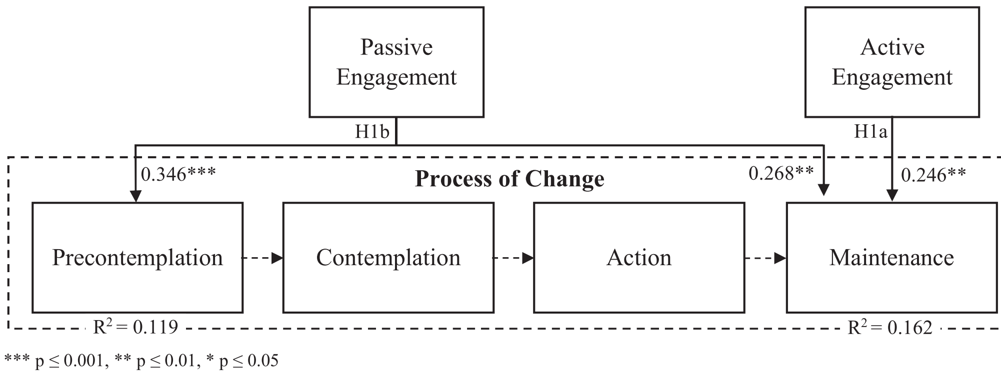


Fig. 3. Influence of active and passive engagement on the stages of change.

significantly and positively engagement. The more motivated users are to participate, the more they actually participate. H2 is supported.

Motivation to participate in r/StopSmoking was found to be significantly influenced (see Table 6) by all uses and gratification factors: information-seeking, status-seeking, providing social support, seeking social support, and entertainment factors. Path coefficients of our model indicate that the motivation to participate in r/StopSmoking has a strong effect on the individuals' active engagement (0.496) and on individuals' passive engagement (0.335), both being highly significant ($p < 0.001$).

3.2.3 Seeking and Providing Social Support Are Motivational Factors That Increase Participation in Online Smoking Cessation Communities (H3). Looking further into the motivation to participate construct enabled us to further test hypotheses H3. The most influential motivational factors are as follows: providing social support (0.340), followed by seeking social support (0.328), information-seeking (0.259), status-seeking (0.251), and entertainment (0.245). Thus, providing and seeking social support are the most influential motivational factors. H3 is supported.

4 STUDY 2: UNDERSTANDING INTERACTIONS ON THE DIGITAL COMMUNITY

Whereas Study 1 built a model based on survey data, linking motivation to participate to engagement and eventually to smoking cessation, Study 2 seeks to assess actual behavior and to provide further insights into the model by looking at the actual engagement on the r/StopSmoking thread by addressing six open questions (see Table 7). To confirm if the identified motivational factors of Study 1 are indeed satisfied, we analyzed the community's interactions, paying particular attention to the kind of information available (Q1) and the proportion of messages aiming to provide or seek social support (Q2). Second, it was discovered in Study 1 that engagement is beneficial to the quitting process, thus we aimed to investigate the dynamics of this engagement in Study 2. Specifically, Study 2 focused on active engagement, both in terms of number of users (Q3) and of interactions, such as posts and comments (Q4).[1] Third, as Study 1 indicated community support for several stages of the process of change, Study 2 aims to understand which stage users are at and the stage targeted by messages (Q5, Q6).

[1]Note that passive participation could not be measured through the available data.

Table 7. Study 2 Questions of Interest

Dimensions	Questions of interest
Motivation to Participate	**Q1:** What type of information is available? **Q2:** How much social support is provided?
Engagement	**Q3:** How has the community evolved in terms of active users? **Q4:** How has the community evolved in terms of active engagement?
Process of Change	**Q5:** What stage of their process are the contributing authors at? **Q6:** Which stage of the smoking cessation process is targeted by the messages?

4.1 Method

To extract the data from r/StopSmoking, the Pushshift API was employed. Pushshift is a big-data storage and analytics project containing the copy of Reddit comments and posts [1]. The use of this API has been chosen over the official Reddit API, as it allows us to export large quantities of Reddit data without quantity limitation. We downloaded 125,349 posts and 803,611 comments resulting in a sample of messages going from the creation of the subreddit to December 31, 2020. As a first step in pre-processing the collected data, deleted posts and comments were removed from the dataset. Deleted posts and comments are still present in Reddit but are easily identifiable as the text message contains a "[deleted]" mention. Orphan comments, i.e., comments belonging to a deleted message, were kept as they could still contain content of interest. The final dataset consisted of 86,554 posts and 745,428 comments written by 92,046 different authors from November 6, 2009 to December 31, 2020.

4.2 Analysis and Results

For each question investigated in Study 2 (Table 7), the sections below present the specific analysis and results obtained.

4.2.1 What Type of Information Is Available (Q1)? To answer Q1, a topic modeling technique was used to identify the key topical interests. The notion of topic has to be understood as a mixture of words used together in similar contexts that potentially allow us to determine the main streams of discussion in r/StopSmoking. As previous studies suggested that the topic of an online discussion is prone to change as the discussion progresses [41], we considered each post and comments distinctly as messages, as we expect many of the longer discussions to have multiple topics. To reduce noise and extract key words from messages, the data were preprocessed to clean and normalize the text messages. The URLs were removed from the messages, and each message was then tokenized. From the list of tokenized words, non-ASCII characters were removed, characters were converted to lowercase, and the punctuation and English stopwords were removed. Stopwords are high-frequency words such as *the*, *to*, and, *also* that are usually filtered out of texts, as they provide little lexical content. In addition to the usual English language stopwords provided by the topwords corpus of the Python Natural Language Toolkit, we added a list of additional stopwords related to the context of smoking cessation. In fact, words such as *quit*, *smoking*, *cigarette*, or *pack* are extensively used in all topics as they are part of the main concern of the community. Excluding such contextual stopwords allowed us to emphasize less common lexical content to widen the gap between the different topics. Finally, a validity check performing a systematic analysis over 5,000 randomly selected messages confirmed our position in the topic modeling. To code topics, we used the **Latent Dirichlet Allocation (LDA)** approach [3], which is one of the most widely used methods to understand the key topics from a large quantity of documents. LDA is an unsupervised algorithm that uses a generative model that uncovers topics by considering the posterior probability of the topics. The lack of a ground truth dataset and the effectiveness LDA demonstrated in previous studies assured us in our choice. Indeed, the analyses of multiple health-related digital communities have demonstrated the effectiveness of LDA topic modeling, such as in the analysis of cannabis consumption influence [30], symptoms and medical usages analysis [44], or general health discussions [40]. Furthermore, using LDA as opposed to most other unsupervised clustering techniques allows us to consider each message with multiple topics. Unfortunately, using LDA requires a predetermined number of topics to be set. However, an interactive topic model visualization tool called pyLDAvis [33] provides help in fitting the LDA models with the optimal number of topics. After fine-tuning this number, the following five topics emerged:

- *Topic 1: Encouragement.* Messages featuring common usage of the terms *congrats*, *congratulations*, *strong*, *proud*, or *easier*. Such *encouragement* messages usually follow the publication of an accomplishment by a community member, as for instance "day 1 wish me luck!" or "3 days without smoking! First two days were horrible, I feel I'm about to relapse." Users easily congrat the accomplishment of other members and provide them with some encouragement such as "Congrats! Stay strong! It will get easier with time!"
- *Topic 2: General information.* Features the words *addiction*, *brain*, or *understand*. These terms seem to refer to authors trying to provide a better understanding of the addiction. Community members share information about what they know or they heard regarding tobacco addiction. For instance, "It's the addiction to smoking that caused you to feel so terrible in the first place. For a decade, you have been feeding your brain dopamine by smoking those poisonous things. Your brain will have a chemical imbalance for some time after quitting. You have to power through that and your brain will rewire itself to produce dopamine normally. Some say it can take 30 to 90 days at most."
- *Topic 3: Personal experience.* Features terms relating to personal experience over smoking cessation process. Words like *weight*, *eat*, *water*, or *exercise* are among the most relevant ones in this cluster. Messages describing personal experience such as "exercise really helped me, gave me purpose and a great feeling 4 times a week."
- *Topic 4: Nicotine substitute.* Most relevant featured terms are *nicotine*, *vape*, *gum*, or *patches*. In this topic, users seem to discuss nicotine substitutes. As it is a hard step to get rid of the nicotine addiction, smokers may share solutions to fight the addiction once they are trying to keep smoke free. A typical message might be "I hated the control smoking had over how I lived my life. I used a vape as a method to quit" or "I've tried quitting a few times over the past few years using patches, nicotine gum, cold turkey, chewing tobacco [...] finally had enough with smoking and started taking Chantix on 4/4."
- *Topic 5: Pieces of Advice.* Reading the Allen Carr's book "The easy way to stop smoking" is among the top advice that community members receive on r/StopSmoking. Topic 5 contains words like *allen*, *carr*, and *way* but also *app*, *audiobook*, or *advice*, and seems to refer to all sorts of advice given when people are willing to quit. For instance, "Allen Carr's book, The Easy Way To Stop Smoking. Read it, believe it, set yourself free from the slavery of nicotine addiction."

4.2.2 How Much Social Support Is Provided (Q2)? To answer Q2, 5,000 randomly selected messages were manually tagged by three researchers in search of "providing social support" and "seeking social support" messages. The 5,000 messages were divided in three for the three researchers, checking inter-rater reliability on 10% of the whole sample. A Cohen's kappa greater than 0.6 was found. On the basis of these 5,000 tagged messages, we trained multi-label classification algorithms to be able to classify messages providing social support and messages seeking social support but also messages doing both (providing and seeking) or neither. Examples of providing social support and seeking social support messages could be, for instance, "Congratulation on your first day

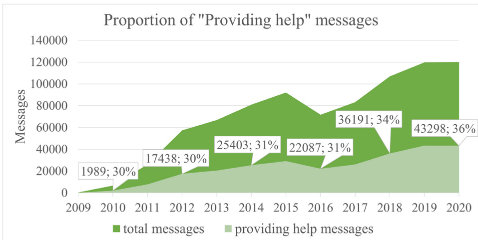


Fig. 4. Proportion of "Providing social support" messages in relation to the total number of messages per year.

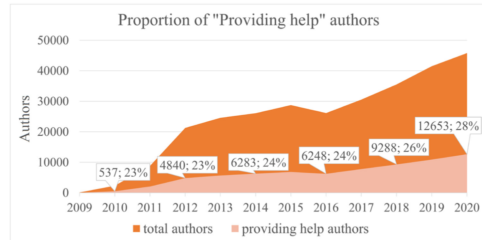


Fig. 5. Proportion of "Providing social support" authors in relation to the total number of contributing authors per year.

Table 8. Confusion Matrix of the "Providing Social Support" Messages Classifier

		Predicted	
		Positive	Negative
Actual	Positive	542	195
	Negative	159	604

without smoking!" and "Wish me luck, I quit smoking today!" Messages were first cleaned and normalized by converting them to lowercase, removing html-tags, punctuation, and non-alphabetic characters. Then stopwords were removed using the same technique as for the clustering under Q1. Next, stemming was applied, transforming words with roughly the same semantics to one standard form. After splitting the dataset into train and test sets, we summarized the messages into numerical vectors using the **Term Frequency Inverse Document Frequency (TF-IDF)** technique. TF-IDF picks the most frequently occurring terms (term frequency or TF) but also measures how unique a word is, i.e., how infrequently the word occurs across all messages (inverse document frequency IDF). We were indeed interested in extracting features corresponding to the terms that frequently occur in the messages belonging to the category we wanted to analyse. Compared to other techniques such as bag of words, the TF-IDF technique is well adapted to extract adapted features for a text classification task, because it solves the problem of less frequent words. Multiple multi-label algorithms were then tested in search of the best accuracy, recall, and precision: One-vs-Rest, Binary Relevance, Classifiers Chains, Label Powerset, and ML-KNN. The One-vs-Rest algorithm provided the best performance and was therefore selected. Classification of messages providing social support provided an accuracy of 0.76, a recall of 0.79, and a precision of 0.76 with the following confusion matrix (Table 8). Classification of seeking social support messages did not provided sufficient performance quality and was then discarded. Figure 4 illustrates the ratio of messages providing social support in respect of the total amount of messages. The results show that the proportion of messages providing social support is slightly growing over time and represents in 2020 more than one third of the total messages. Figure 5 also shows the proportion of authors providing social support among total number of contributing authors. From 2010 to 2020, the proportion of authors providing social support increased from 23% to 28%.

4.2.3 How Has the Community Evolved in Terms of Active Users (Q3)? In the year 2020, 45,819 distinct authors posted a message on r/StopSmoking, from which 15,720 were new on that thread. To have a better idea of the community evolution, the annual number of contributing authors and of first-time contributing authors have been extracted. This allowed us to identify four main stages

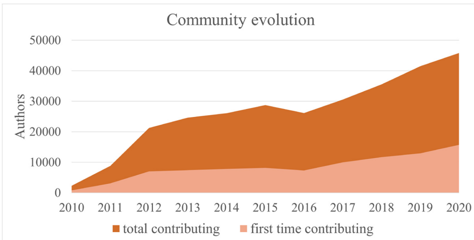


Fig. 6. Community evolution in term of contributing authors and first-time contributing authors.

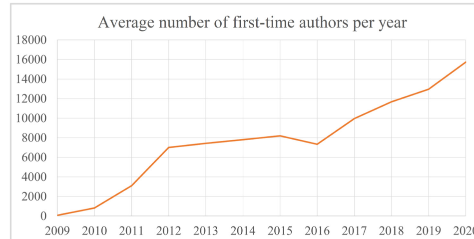


Fig. 7. Average number of first-time contributing authors per year.

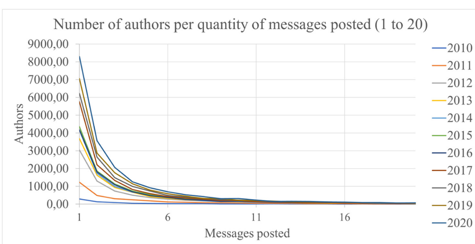


Fig. 8. Comparison by years of the number of authors having posted the same quantity of messages.

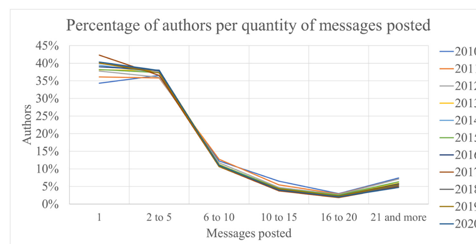


Fig. 9. Comparison of the percentage of authors per year having posted the same amount of messages.

of evolution for the r/StopSmoking community: (1) From 2010 to 2012, the number of authors and new contributing authors exploded, with, for instance, an increase of 279.5% first-time contributing authors from 2010 to 2011. (2) From 2013 to 2015, the annual growth rate of the community appears to have stabilized. For that period, the mean growth rate of the number of authors contributing was 10.6% with a 4.9% **standard deviation (SD)**, while the mean growth rate of the number of first-time authors was 5.4% with a 0.6% standard deviation. (3) In 2016, there is a growth rupture. Topic extraction could not reveal anything different for that specific year and this decreased growth remains at this point unexplained. (4) From 2017 to 2020, a stable period in terms of growth starts again. In fact, the weekly mean growth rate from 2017 to 2020 is 15.1% with a standard deviation of 3.1% for overall contributing authors and 21.4% with a 10.6% standard deviation for the first-time contributing authors. Figure 6 allows us to visualize these four stages. To have a better idea of the author rollover, yearly means of first-time contributing authors have been extracted. Figure 7 presents the average number of first-time authors per year. Another interesting insight concerning community authors comes from Figures 8 and 9, which show that a greater part of the authors contribute only a few times in r/StopSmoking, with a large part contributing only once.

4.2.4 How Has the Community Evolved in Terms of Active Engagement (Q4)? During 2020, the r/StopSmoking community registered 12,344 new posts and 107,663 new comments, with an average of 233 posts and 2,031 comments per week. Figures 10 and 11 present the evolution of such figures from the beginnings of r/StopSmoking until end of 2020. The evolution of the amount of posts and comments registered every year appears to follow the same four-stage pattern of the community evolution presented in Q3. To have a better idea of the frequency of participation, we extracted weekly figures. The weekly mean number of posts is illustrated in Figure 12, the weekly mean number of comments in Figure 13, and the weekly mean number of contributing authors in Figure 14. Regarding the votes that a message on r/StopSmoking can receive, Figure 15 illustrates

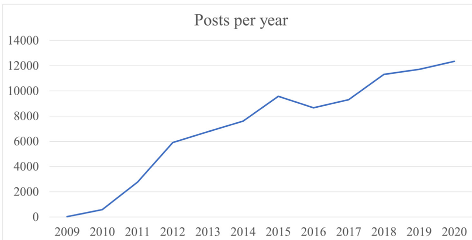


Fig. 10. Total number of posts per year.

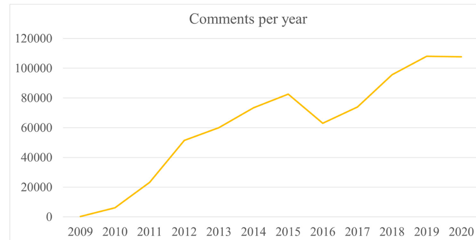


Fig. 11. Total number of comments per year.

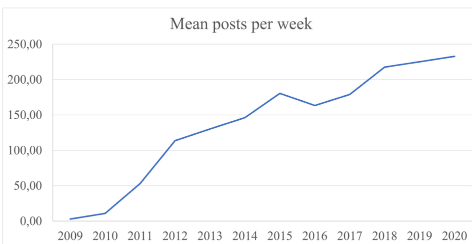


Fig. 12. Average weekly posts year on year.

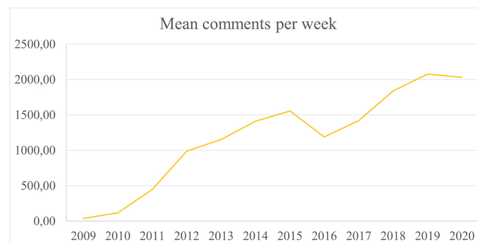


Fig. 13. Average weekly comments year on year.

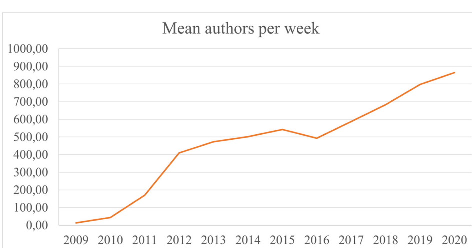


Fig. 14. Average weekly number of contributing authors year on year.

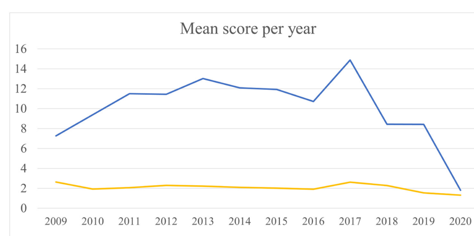


Fig. 15. Mean messages score per year.

the mean messages score per year, the score of each message represents the addition of upvotes with the subtraction of downvotes, that are voluntarily assigned by the community. The comments scores appears to be stable through the years, while the posts score seems be more subject to variation, with a minimum reached in 2020.

To investigate the possible activity cycles, we plotted the average number of messages posted on r/StopSmoking week by week (Figure 16). We could see that the first weeks of years 2012, 2013, 2015, 2016, 2017, 2018, 2019, and 2020 are marked by a peak in activity. After manually analyzing the messages, we could attribute these peaks of activity to the New Year's resolutions—smokers deciding to quit smoking for the new year and then seeking socialization, information, status, or entertainment on r/StopSmoking. Further analysis of the community attendance allowed us to determine that the community is more active during the working days (Figure 17), particularly Tuesday and Wednesday. Peak hours of activity are from 4 pm to 8 pm (UTC) (Figure 18).

4.2.5 What Stage of Their Process Are the Contributing Authors at (Q5)? To answer Q5, the 5,000 randomly selected messages were manually tagged by three researchers whether they were written by smokers, by former smokers, or if it was not specified or deducible. For instance, an example

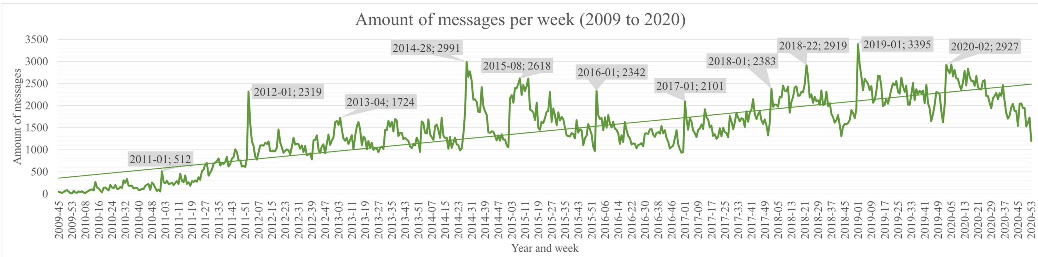
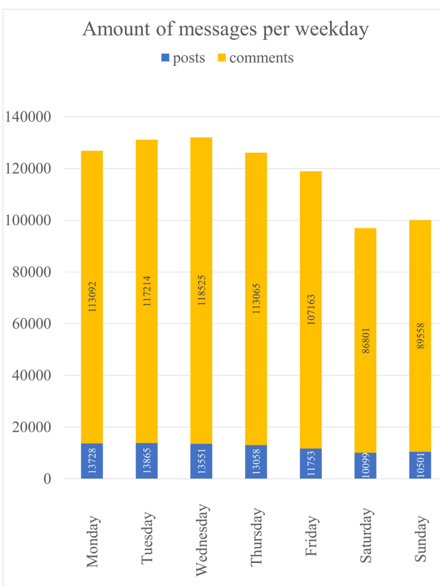


Fig. 16. Overall number of posts and comments per week day.



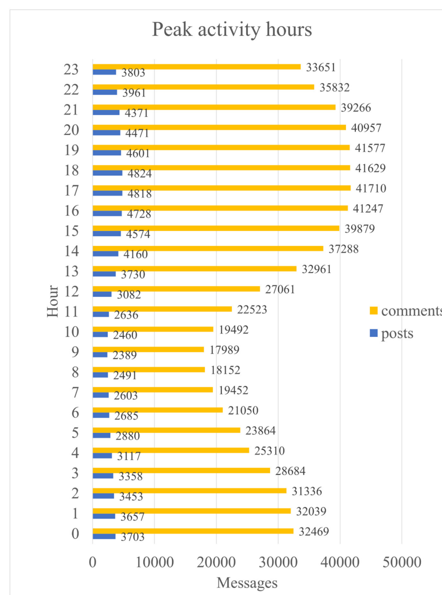


Fig. 17. Overall number of posts and comments per week day.

Fig. 18. Overall number of posts and comments per hour (UTC).

of a smoker's message would be "Saw all the class stories on here about people quitting and was looking for some tips. I've wanted to quit smoking for a long time, but always fall back into it," while an example of former smoker's message would be "I've also been smoke-free for 1 month today. What can we expect in terms of benefits of quitting going forward?" The 5,000 messages were divided into three for the three researchers checking inter-rater reliability on 10% of the whole sample. A Cohen's kappa greater than 0.65 was found. On the basis of these 5,000 tagged messages, we trained a classification algorithm to be able to classify messages according to the status of its author. Messages were first cleaned and normalized by converting them to lowercase and removing html-tags, punctuation, and non-alphabetic characters. Then stopwords were removed using the same technique as for the clustering under Q1. Stemming was then applied, transforming words with roughly the same semantics to one standard form. After splitting the dataset into train and test sets, we summarized the messages into numerical vectors using TF-IDF technique. The One-vs-Rest algorithm was used to classify the messages. Classification of messages authored by current smokers did not provide sufficient performance quality and were then discarded. Classification of

Table 9. Confusion Matrix for Messages Authored by Former Smokers Classifier

		Predicted	
		Positive	Negative
Actual	Positive	868	164
	Negative	178	290

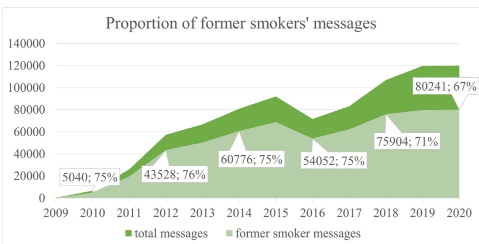


Fig. 19. Proportion of former smokers messages in relation to the total number of messages per year.

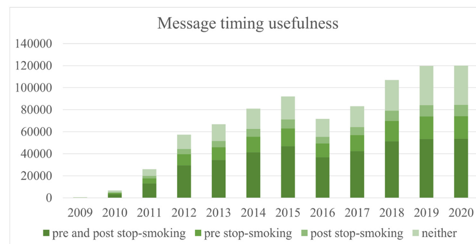


Fig. 20. Estimation of the message timing usefulness year by year.

messages authored by former smokers, provided an accuracy of 0.77, a recall of 0.62, and a precision of 0.64 with the following confusion matrix (Table 9). As classifying former smokers messages performed better, we applied this classifier on the whole dataset obtaining results illustrated in Figure 19. The results of our analysis indicates that 67% of messages posted in 2020 were authored by users classified as former smokers, which suggests that much of the activity on this community is conducted by individuals in the maintenance stage of their TTM process of change.

4.2.6 Which Stage of the Smoking Cessation Process Is Targeted by the Messages (Q6)? The 5,000 randomly selected messages were manually tagged by three researchers looking for timing usefulness messages. Timing usefulness was classified as "pre stop-smoking" if the message was useful before the action stage of the TTM process of change and as "post stop-smoking" if it was useful after the action stage of the TTM process of change. For instance, a message useful before the action of quitting could be "Just say no, start doing something else to take your mind off it. You need to read Allen Carr's book." A message useful after the action of quitting could be "Sometimes, when I'm very stressed (it's been a stressful week at work), or when I see someone smoking, I feel a strong craving, but I remind myself that I've been good for 18 days, and I don't want to throw it all away." The 5,000 messages were divided in three for the three researchers checking inter-rater reliability on 10% of the whole sample. A moderate inter-rater reliability was found among researchers with a 0.45 Cohen's Kappa. On the basis of these 5,000 tagged messages, we trained multi-label classification algorithms to be able to classify messages on the basis of their timing usefulness on the smoking cessation process of change. Messages could contain information or advice that is useful not only before quitting (pre stop-smoking) and after quitting (post stop-smoking) but also for both situations or neither of them. Multi-label algorithms were then tested in search of the best accuracy, recall, and precision. The One-vs-Rest algorithm was selected as it provided the best performance compared to Binary Relevance, Classifiers Chains, Label Powerset, and ML-KNN. Messages were preprocessed by lowercasing them and removing html-tags, punctuation, and non-alphabetic characters. Then stopwords were removed using the same technique as for the clustering under Q1. Next, stemming was applied, transforming words with roughly the same semantics to one standard form. After splitting the dataset into the

Table 10. Confusion Matrix of the Pre Stop-smoking Messages

		Predicted	
		Positive	Negative
Actual	Positive	917	134
	Negative	229	220

Table 11. Confusion Matrix of the Post Stop-smoking Messages

		Predicted	
		Positive	Negative
Actual	Positive	835	170
	Negative	207	288

train and test sets, we summarized the messages into numerical vectors using TF-IDF technique. Classification of messages useful before action provided an accuracy of 0.76, a recall of 0.50, and a precision of 0.62 with the following confusion matrix (Table 10). Classification of messages useful after action provided an accuracy of 0.75, a recall of 0.58, and a precision of 0.63 with the following confusion matrix (Table 11). Figure 20 illustrates the proportion of messages being useful before having stopped smoking, after having stopped smoking, and in both or neither situation. It appears that a majority of messages are useful in both pre and post stop-smoking situations. Comparing messages that are exclusively useful before or after having stopped smoking, figures suggest that messages being useful to people before they stop smoking are increasingly present year on year.

5 STUDY 3: UNDERSTANDING THE IMPACT OF COVID-19 ON THE DIGITAL COMMUNITY

Study 3 aims at understanding how interactions on r/StopSmoking were affected by a special crisis, i.e., the COVID-19 pandemic. In 2020, the COVID-19 pandemic shook the whole world by killing several million people around the globe and affecting billions indirectly through economic hardship or politically imposed lockdowns. As a result, stress levels and conditions like anxiety and depression seem to have have increased [27, 43]. Furthermore, with social distancing, highly effective support relying on a face-to-face setting will probably have lower participation rates than usual [38]. Because of this, it has been suggested that social media could fulfill a compensatory function by substituting physical contact with a virtual touch and extending social contacts beyond the physical boundaries of COVID-19 confinement [2]. These findings and predictions about the impact of the COVID-19 crisis in general led to the following hypotheses regarding its impact on the r/StopSmoking community in particular:

- **H4a:** During the COVID-19 crisis, stress levels have increased among r/StopSmoking users.
- **H4b:** During the COVID-19 crisis, anxiety levels have increased among r/StopSmoking users.
- **H4c:** During the COVID-19 crisis, depression levels have increased among r/StopSmoking users.
- **H5:** r/StopSmoking engagement has increased during the COVID-19 crisis.
- **H6:** Perceived usefulness of r/StopSmoking has increased during the COVID-19 crisis.

To do this analysis, Study 3 takes a multimodal analytics approach combining survey results and activity traces of r/StopSmoking interactions during the period 2017 to 2020.

5.1 Method

We designed a survey to assess people's perception of their own levels of stress, anxiety, and depression as well as the usefulness of r/StopSmoking both *before* and *during* the COVID-19 crisis. In this survey, we first probed the community's general concern and feeling about the COVID-19

Table 12. Perceived Usefulness, Stress, Anxiety, and Depression before and during COVID-19 Crisis

	Before COVID-19		During COVID-19		Significance
	Mean	**SD**	**Mean**	**SD**	**Two-tailed *t*-test**
PUS-6 Score	3.71	0.72	3.88	0.75	1.50 (not significant)
PSS-4 Score	2.74	0.73	3.01	0.77	2.64 (significant at $p < 0.05$)
STAI-6 Score	3.11	0.55	3.21	0.67	1.02 (not significant)
PHQ-2 Score	2.02	0.74	2.46	0.84	3.54 (significant at $p < 0.01$)

crisis. Then r/StopSmoking user engagement before and during the COVID-19 crisis was assessed. Engagement was measured by asking participants how frequently they performed the following types of activities: visiting, reading, posting, commenting, and voting. Next, respondents evaluated the perceived usefulness of r/StopSmoking as well as the perceived stress, anxiety, and depression levels during and, if applicable, before the COVID-19 crisis. The "before" section was only displayed if the respondents confirmed their having engaged in r/StopSmoking before the COVID-19 crisis. In that case, before and during subsections were randomized. To measure the perceived usefulness of r/StopSmoking, we used the validated Perceived Usefulness Scale from the work of David et al. [12]. The self-reported perceived stress of respondents was measured thanks to the PSS [10] in its short form, PSS-4, as validated by the study of Warttig et al. [59]. To monitor self-reported perceived depression, we used the PHQ in its two-item form (PHQ-2) as suggested and validated by Löwe et al. [32]. To self-report perceived anxiety, we used one of the most frequently used measures of anxiety in applied psychology research, the Spielberger STAI [55] (in its short form, STAI-6), as presented and validated by Marteau et al. [35]. All of these validated scales were adapted to the context. As with the first survey, it was validated by the University of Neuchâtel ethics committee, informed consent was requested, and participants were free to stop at any time. Users of the r/StopSmoking subreddit were invited to participate in the survey through links posted directly to the site. The survey was completed on a voluntary basis, and no compensation was given to respondents. A total of 211 responses were collected from 20 January 20 to March 21, 2021. To maintain the visibility of the invitation among other posts, invitations to participate in the survey were randomly re-posted throughout the data collection period. After preliminary analyses and data preparation for 57 respondents were retained, as they were engaged in r/StopSmoking before and during COVID-19 crisis; most of the other respondents did not engage in r/StopSmoking before the crisis.

5.2 Analysis and Results

For each hypothesis investigated in Study 3, the sections below present the specific analysis and results obtained.

5.2.1 During the COVID-19 Crisis, Stress Levels Have Increased among r/StopSmoking Users (H4a). The four-item instrument asks respondents to rate how often they experienced stressful situations on a Likert scale ranging from 0 to 5, where 0 = never and 5 = very often. Two of the PSS-4 items are reverse scored, and so these variables were recoded. Higher values on the PSS-4 indicate more stress. As presented in Table 12, on a maximum score of 5, the perceived stress of our sample increased from 2.74 to 3.01 with standard deviations of respectively 0.73 and 0.77 representing an increase of 10%. Looking more in detail at respondents' answers, we noted that PSS-4 Score increased for 58% of our sample. A two-tailed repeated-measures *t*-test showed that r/StopSmoking users perceive themselves significantly ($p < 0.05$) more stressed during the pandemic. H4a is supported.

5.2.2 During the COVID-19 Crisis, Anxiety Levels Have Increased among r/StopSmoking Users (H4b). The six-item measure asks respondents to rate how they felt on a Likert scale going from 0 to 5, where 0 = not at all and 5 = very much so. Respondents had to rate whether they felt calm, tense, upset, relaxed, content, or worried. Three of the STAI-6 items are reverse scored, and so these variables were recoded. Higher values on the STAI-6 indicate more anxiety. As shown in Table 12, on a maximum score of 5, the perceived anxiety of our sample increased from 3.11 to 3.21 with standard deviations of respectively 0.55 and 0.67 representing an increase of 3.1%. Looking in more detail, STAI-6 Score increased for 46% of our sample. However, a two-tailed repeated-measures t-test failed to show a significant difference in perceived anxiety. H4b is not supported.

5.2.3 During the COVID-19 Crisis, Depression Levels Have Increased among r/StopSmoking Users (H4c). The two-item questionnaire, PHQ-2, asks respondents how often they were affected by (1) having little interest or pleasure in doing things or (2) feeling down, depressed, or hopeless on a Likert scale going from 0 to 4 where 0 = not at all and 4 = nearly every day. Higher values on the PHQ-2 indicate higher level of depression. As shown in Table 12, on a maximum score of 4, the perceived depression of our sample increased from 2.02 to 2.46 with standard deviations of respectively 0.74 and 0.84 representing an increase of 22%. More specifically, PHQ-2 Score increased for 49% of respondents. A two-tailed repeated-measures t-test showed that r/StopSmoking users appear to perceive themselves significantly ($p < 0.05$) more depressed during the pandemic. H4c is supported.

5.2.4 r/StopSmoking Engagement Has Increased During COVID-19 Crisis (H5). To verify this effect, we relied directly on data retrieved during Study 2. We defined the COVID-19 period of relevance for r/StopSmoking from January 24, 2020 to December 31, 2020. These dates were chosen because the first time that a user of r/StopSmoking talked about the COVID-19 disease was on week 4 of 2020 (January 24, 2020): "Another reason to quit - Wuhan Virus - This virus puts a great toll on the respiratory system - lung function is very important," and because the data collected stopped on December 31, 2020. Then we compared the COVID-19 period with the past ordinary evolution of the r/StopSmoking community. But the community has a global yearly natural growth trend and eventual seasonality, which should be taken into consideration. Study 2 suggested in fact a seasonal effect, with peaks of activity in the early weeks of each year (probably due to New Year's resolutions).

Therefore, we made predictions on the COVID-19 period, based on the previous growth trend and seasonality. To establish predictions, we focused on data from 2017 to 2019, as, according to Study 2, they belong to the same stage of evolution of the community. The general evolution trend of users' interactions was extracted by identifying the regression line, while seasonality was added by calculating the seasonal index of every week [23]. For instance, in Table 13, we present the weekly number of posts recorded in r/StopSmoking for years 2017 to 2019. On that basis, the mean for each week for years 2017 to 2019 was calculated. Then with the 52 weekly means the overall total mean (207.3) was calculated. Finally, the seasonal index of each week is calculated by dividing the week mean by the total mean (e.g., for Week 1 293.3/207.3 = 1.4).

Predicted mean number of posts and comments for COVID-19 period were calculated, making use of the regression line and multiplying each week estimation by the seasonal index of the corresponding week. To check if there was a statistically significant difference between calculated and measured values, we performed a bilateral Student's test of the residuals with a hypothetical mean of zero. The residuals matrices of both posts and comments were normally distributed

Table 13. Seasonality Calculus

	Week 1	*Week 2*	*Week 3*	...	*Week 12*	*Week 13*	*Week 14*	...	*Week 50*	*Week 51*	*Week 52*	*total mean*
2017	245	189	173	...	148	172	169	...	154	142	160	
2018	307	213	229	...	185	215	232	...	203	189	190	
2019	328	268	270	...	217	249	261	...	221	222	208	
mean	293.3	223.3	224.0	...	183.3	212.0	220.7	...	192.7	184.3	186.0	207.3
season coeff.	1.4	1.1	1.1	...	0.9	1.0	1.1	...	0.9	0.9	0.9	

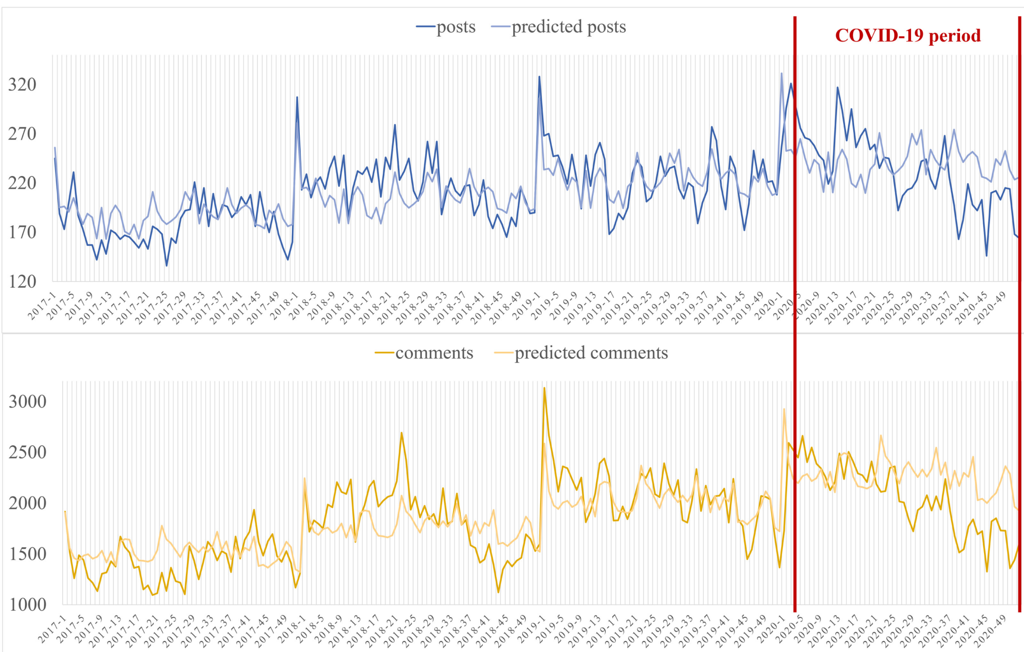


Fig. 21. Predicted vs. effective interactions. On the top number of posts per week and on the bottom number of comments per week.

(Shapiro-Wilk test, $p > 0.1$). The difference between measured and calculated numbers of comments was significantly different, but that was not the case for the number of posts. The actual numbers of comments per week during 2020 (M = 2033.0, SD = 334.0) demonstrated a significantly lower number of comments from what could have been calculated (M = 2262.0, SD = 153.0), $t(47) = 5.06$, $p < 0.001$. The results suggest that the differnce observed between the predicted vs. actual number of comments may be due to the special situation of the crisis. H5 is partially supported (more comments but not more posts). Looking at the curves more closely (Figure 21), we could observe a sudden drop of the number of comments per week with two negative peaks on Week 29 (July 13–19, 2020) and 39 (September 21–27, 2020). Regarding the posts, we can still observe an interesting second peak of posts at the beginning of the COVID-19 period that could not be predicted by the prediction curve.

5.2.5 Perceived Usefulness of r/StopSmoking Has Increased during COVID-19 Crisis (H6). The six-item instrument measuring the perceived usefulness (PUS-6) of r/StopSmoking was measured on a Likert scale going from 0 to 5 where 0 = strongly disagree and 5 = strongly agree. Questions measured whether r/StopSmoking made it easier to get support, improved willpower, increased the motivation, was useful in supporting them in their smoking cessation process, and/or helped in making a quit attempt. Higher values on the PUS-6 indicate a higher level of perceived usefulness. As shown in Table 12, on a maximum score of 5, the perceived usefulness of our sample increased from 3.71 to 3.88 with standard deviations of respectively 0.74 and 0.84 representing an increase of 4.5%. More specifically, PUS-6 Score increased for 51% of respondents. However, according to a two-tailed repeated-measures this perceived usefulness growth is not significant. H6 is not supported.

6 DISCUSSION

Our overarching findings obtained through Studies 1, 2, and 3 provide support that engagement in smoking cessation communities is positively correlated to the process of change (H1). This means that the frequency of participation in r/StopSmoking has an influence on the overall process of change. The results show a significant but modest relation between the overall engagement and the overall process of change.

Engagement with such communities help smokers at the beginning of their process and also in the maintenance of their withdrawal. Therefore, encouraging smokers to engage in online smoking cessation communities should be considered as a good practice. In fact, such engagement could in a first instance increase their awareness of the unhealthy behavior, enable them to realize the negative impacts, and feel fear about it but also feel hope when they hear about how people can change to healthy behaviors. These processes can potentially lead to a first success: a firm intention to quit smoking in the next 6 months. Engaging smokers who have already taken action and are trying to stay smoke free has also to be considered as a good practice, as this research provides further evidence that such engagement helps in staying smoke free. In fact, Study 1 indicates that active engagement is significantly correlated to the maintenance stage, which means, for instance, that the frequency of commenting or voting helps ex-smokers to not relapse. Active engagement allows us to obtain social support from peers, but we might expect that it would also encourage people to provide social support in turn to peers in a lower level of accomplishment in their process of change. An additional line of inquiry might look at a possible feedback loop between such engagement and motivation to participate.

Looking at the user interactions (Study 2), we validate the consistency of results of Study 1 obtained through surveys by looking at actual behavior of the online community. Analyzing the conversations, we can observe that most of the active users engaged in r/StopSmoking have already stopped smoking (Q5) and are mainly sharing messages helping others to stop smoking and stay smoke free (Q6). About one third of authors and one third of messages are aimed at providing social support (Q2). This result further validates what was found in Study 1 on the motivational factors side. Providing social support is the main motivational factor to participate in r/StopSmoking (H3), and it represents one third of r/StopSmoking content. This result supports the presence of *expert patients* who are willing to provide experience and support to newcomers, thereby helping themselves to stay smoke free [17]. By helping others, former smokers are helping themselves to stay smoke free. Former smokers should therefore also be encouraged to use such online smoking cessation communities.

Further results on motivational factors show that all uses and gratification motivational factors are positively linked to engagement in online smoking cessation communities (H2). Information-seeking, status-seeking, socialization (providing and seeking social support), and entertainment are all significantly and positively contributing to the motivation to participate and consequently to the online engagement of the user.

A deeper analysis of the community interactions in Study 2 shows that r/StopSmoking keeps growing, with increasing numbers of weekly messages (Q4) and authors (Q3). More and more smokers are therefore helped. Even if most users seem to only seldomly actively engage in r/StopSmoking (Q3), the increasing quantity of messages generated every day provides ever-new material feeding individuals willing to get passively or actively engaged.

When analyzing more closely the content of such interactions, we can see that the main topics of conversation in r/StopSmoking (Q1) are as follows: encouragement, general information, personal experience, nicotine substitute, and pieces of advice. These topics fit surprisingly well with strong efficacy evidence-based interventions for smoking cessation (brief advice, behavioral support, pharmacotherapy, and abstinence evaluation) [58]. Brief advice can be found in general information, personal experience, and pieces of advice. Behavioral support is potentially provided through encouragement and pieces of advice. Pharmacotherapy is discussed within the topic of nicotine substitute and personal experience. Abstinence evaluation is also frequently self-evaluated through relating personal experience, as one of the most common posts relates to the number of days since the author stopped smoking.

These insights provided by the core content of the community could also partly explain some of the reason behind the finding established in Study 1 showing that the more frequently someone visits and reads content on r/StopSmoking the more likely they will be to express a firm intention to quit smoking or stay smoke free if they have already quit (H1b). The impact of the participants' engagement with topics supported by strong efficacy evidence-based interventions could potentially provide more insight into the understanding of the smokers' progress in their process of change.

Finally, Study 3 investigates how the community reacted to a global crisis, namely the COVID-19 pandemic. Our results confirm that users seem to have increased stress levels (H4a) and increased signs of depression (H4c), but that was not true for anxiety (H4b).

Furthermore, Study 3 hints at the fact that the community's perceived usefulness did not increase during the crisis (H6) and engagement even dropped (H5). The number of posts did not present a significant difference from what could have been predicted, but the number of comments did: Fewer comments than predicted were found. This hints at the fact that the need for support has potentially increased in the pandemic, but the r/StopSmoking community has not been able to fully respond to these needs. Perhaps this decrease of comments is due to lack of time or a reordering of people's concerns [2]. This may mean that the number of people visiting the community to express themselves spontaneously has not significantly changed. However, the number of people reacting to these messages has decreased. One might wonder if the online community has lost its usefulness, but according to our results there has been no significant change in the perceived usefulness of the online community. Furthermore, semantic analysis of messages posted during the COVID-19 period revealed that users who continued to use the community took COVID-19 mostly as an opportunity to stop smoking. This latter result may be in line with a parallel research investigating the effect of the COVID-19 pandemic on smoking cessation success for patients admitted to clinics and those supported by phone [26]. This support provided remotely, whether by phone or via the online community, may have been used as a way to vent excess stress and depression during this period and although the community was less responsive, it was still perceived as useful. A more detailed and longer-term analysis would certainly shed more light on this change in dynamics and also determine whether it has persisted over time.

Future work could also further investigate how design features of such communities could be adapted to better cope with such situations and avoid such engagement decrease.

6.1 Scientific Contributions

As a contribution to research, our work is a first step into a more nuanced understanding of motivation to participate in digital smoking cessation communities and its behavioral impact on individuals' online engagement and offline smoking cessation process of change. We were able to demonstrate the validity of a model measuring the influence of engagement in online smoking cessation communities on offline smoking behavior as well as potential motivational antecedents to such engagement. TTM's stages of change and its mediating processes have been used to measure the progress between the various stages, showing the significant support of online communities in the precontemplation to contemplation transition as well as on the maintenance to termination transition.

The uses and gratification approach allowed us to identify motivational factors of such digital communities. This research confirmed status-seeking and entertainment as relevant motivational factors of online engagement, these findings being consistent with previous literature [36, 42]. This research also found the information-seeking factor to be one of the most salient motivational factors examined, contradicting previous research on Reddit engagement [36] but being aligned with conclusions of other literature relating to Facebook online communities [48]. With Reddit containing such a variety of topics and therefore of different communities, we believe that uses and gratification can vary from one community to another. Especially when participating in a subreddit, such as r/StopSmoking, where the community is driven by a common health-related goal, the uses and gratifications may be different than another random community driven by a completely different goal. This research extended the uses and gratification approach by extending the socialization factor, giving a better understanding of social interaction motivation. This novel definition of the socialization factor, through providing and seeking social support subfactors, found them to be the most significant motivational predictors of the overall motivation to engage in r/StopSmoking.

This research also provides a multimodal analytics methodology combining survey results and activity traces over an extended period of time. Such analysis has demonstrated that users' motivational factors appear to be aligned with their acts, because we found that one third of the community has basically already stopped smoking and is providing social support to others. The r/StopSmoking community keeps growing in the number of users and the quantity of messages, but the information access infrastructure, i.e., the forum-style design of the Reddit platform, is only marginally adapted to the size and the context of this community, perhaps keeping out a potentially larger audience. For instance, there are no facilities to quickly find adequate messages, advice, or other information personalized to one's situation and context. Insights revealed from this Study 2, i.e., a majority of former smokers among active users, may highlight a lack of design to attract a larger panel of smokers that is more representative of the TTM process of change in its entirety.

This research also offers a better understanding of the evolution of the r/StopSmoking online community through the COVID-19 crisis from different perspectives. For instance, our findings indicate that perceived levels of stress and depression increased among community users, which is in line with concerns raised by early COVID-related studies concerning smokers [27, 43]. However, smokers' perceived anxiety or perceived usefulness of the community have not significantly increased with the COVID-19 crisis, which is less in line with what we might have expected [2, 43]. In terms of activity on the platform, Figure 21 shows an increase of posts that coincides with the start of the first wave (March 2020), and this may indicate that at least at the start of the pandemic, people flocked to the community. However, there is a declining trend of posts and messages over the year 2020. These findings are in line with others who have found the COVID-19 pandemic to have been associated with information overload [28].

6.2 Implications for Practice

In the various processes mediating the progress between stages of change, we observed that early and late stages of the overall process were significantly supported by online engagement. More specifically, we observed highly significant correlation between passive engagement and the stages of precontemplation and maintenance. This means that visiting and reading online communities could give smokers the intent to stop smoking within the next 6 months or help them to stay smoke free when they are not 100% confident about being able to do so. High significance correlation between the active engagement and maintenance stages further confirms the effectiveness of online communities in keeping ex-smokers away from temptation and relapses.

Encouraging smokers to share their experience on online smoking cessation communities could help them in their own process of change. It could also help them to potentially become expert patients who will then be even more helped [17]. Interestingly, the three most relevant factors of motivation to engage in digital smoking cessation communities are aligned with primary care interventions of strong efficacy [58]. Providing and seeking social support on online communities could be assimilated to peer-group behavioral support, while information-seeking could be assimilated to self-help material. Introducing smokers to these digital smoking cessation communities could trigger motivational factors, which in turn would be effective in the process of behavior change. Further research on a long-term basis is still needed to verify these assertions. User interaction data confirmed the potentially effective help that such communities could provide to smokers, as the topics of discussions are interestingly similar to the strong efficacy smoking cessation interventions as presented in primary care guidelines [58]. This is especially the case when the traditional support vectors cannot be reached for whatever reason, as is currently the case within the COVID-19 crisis. Related to COVID, we agree with recommendations made by the previous literature on this subject [27, 43]. Special attention must be given to smokers and former smokers during and after this period, not only because they present an increased risk but also because we have been able to show that potential risks of an increased level of stress and depression exist.

6.3 Limitations and Future Research

This research is not without limitation. A first limitation is that the model only includes a partial view of the smoking behavior of a user in a limited context (e.g., health conditions, demographic context, family context). A second limitation is the fact that we did not integrate the preparation stage in our research model, because there is no clear process allowing people to move from preparation to action. This transition could warrant more attention in future work, since it can be particularly challenging. A third limitation is the scope of the data, which unfortunately does not allow access to the identities of those users who engage more passively in the community. Neither the Reddit API nor the Pushshift API provide access to data such as the identity of users visiting, reading, or even voting on a community message. A fourth limitation lies in the various classifier algorithms that can surely be further improved by including a larger quantity of features and tagged data. We could also extend the analysis of the Reddit messages with the extraction of semantic features and their polarity (i.e., sentiment analysis). A fifth limitation is related to the sample sizes of our surveys, which might not contain representative members of the community. This is particularly true for Study 3, where we could only find 57 participants that engaged with the community both before and during COVID-19 crisis.

Future work could investigate how to encourage passive users to become active and expert patients to become even more active online. Visualizing activity traces could be one way to motivate this behavior [25]. Finally, another interesting research topic would be to test if the model can be applied to other addictions or other behavioral problems (e.g. eating disorders).

7 CONCLUSION

In this research article, we presented a novel model for investigating the influence of digital smoking cessation communities engagement on the behavior of actual smokers in their process of change. We also studied the motivational factors of such participation in online communities while extending the uses and gratification approach to fit the special context of smoking cessation. To do so, we undertook three distinct studies. The first, on the basis of a survey data from 169 Reddit contributors of the r/StopSmoking thread. The second, on the basis of 10 years of user interaction data. The third, on the basis of a survey data from 57 Reddit users contributing to this thread from at least one year as they contributed before and during actual COVID-19 crisis. The results show that active and passive engagement in online smoking cessation communities has a significant influence on the process of change, mainly in the precontemplation and maintenance stages of the TTM stages of change. We identified that uses and gratification motivational factors are correlated to online smoking cessation communities engagement. The novel structure of the socialization factor, including providing and seeking social support, is relevant, because they were found to be the main predictors of the individual's motivation to engage in such virtual communities. Indeed, providing social support was found to be a main motivational factor for contributing to the community, and the analysis of the user interaction data showed that a third of messages are indeed aimed at helping others. We also found that the main topics discussed in the community were tightly linked to strong efficacy evidence-based literature on smoking cessation, i.e., advice, information, pharmacotherapy, and peer support. A broader analysis of the community shows that it is expanding as more users join, and we believe that such communities could be of great help to smokers and former smokers in general, especially in crisis situations, as recently with that of COVID-19. Nevertheless, our results highlight that participation during such a crisis is not to be taken for granted, as participation in the community dropped during that time, and future work should investigate how to better support users in such situations.

Finally, our study showed that engagement in digital smoking cessation communities such as r/StopSmoking is likely to increase in the future, even though it was slowed down by the crisis. Many users are drawn to such communities not only to seek help but also to provide help to peers. Our results indicate that this engagement can lead to a positive impact on people's smoking cessation journey.

REFERENCES

[1] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. 830–839.

[2] Natalie Bazarova. 2020. Staying Connected Via Social Media in the Age of COVID-19. Retrieved from https://www.psychologytoday.com/us/blog/social-media-stories/202003/staying-connected-social-media-in-the-age-covid-19.

[3] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 3 (2003), 993–1022.

[4] Jay G. Blumler and Elihu Katz. 1974. The uses of mass communications: Current perspectives on gratifications research. *Sage Annu. Rev. Commun. Res.* (1974).

[5] Kenneth A. Bollen. 2002. Latent variables in psychology and the social sciences. *Annu. Rev. Psychol.* 53 (2002), 605–634.

[6] Nicholas A. Christakis. 2008. Health care in a web. *Br. Med. J.* 336 (2008), 1468–1468.

[7] Nathan K. Cobb, Amanda L. Graham, M. Justin Byron, David B. Abrams, et al. 2011. Online social networks and smoking cessation: A scientific research agenda. *J. Med. Internet Res.* 13 (2011), e119.

[8] Nathan K. Cobb, Megan A. Jacobs, Jessie Saul, E. Paul Wileyto, and Amanda L. Graham. 2014. Diffusion of an evidence-based smoking cessation intervention through Facebook: A randomised controlled trial study protocol. *Br. Med. J. Open* 4, 1 (2014), e004089.

[9] Sheldon Cohen and Edward P. Lemay. 2007. Why would social networks be linked to affect and health practices? *Health Psychol.* 26, 4 (2007), 410–417.

[10] Sheldon Cohen and Edward Lichtenstein. 1990. Perceived stress, quitting smoking, and smoking relapse. *Health Psychol.* 9, 4 (1990), 466.

[11] Sarah L. Cutrona, Douglas W. Roblin, Joann L. Wagner, Bridget Gaglio, Andrew E. Williams, Rosalie Torres Stone, Terry S. Field, and Kathleen M. Mazor. 2013. Adult willingness to use email and social media for peer-to-peer cancer screening communication: Quantitative interview study. *JMIR Res. Protocols* 2, 2 (2013), e52.

[12] Fred D. Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart.: Manage. Inf. Syst.* 13, 3 (1989), 319–339. https://doi.org/10.2307/249008

[13] Alessio De Santo and Adrian Holzer. 2019. Assessing digital support for smoking cessation. In *Proceedings of the European, Mediterranean, and Middle Eastern Conference on Information Systems*. 331–343.
[14] Arvind Diddi and Robert LaRose. 2006. Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an Internet environment. *J. Broadcast. Electr. Media* 50, 2 (2006), 193–210.
[15] Mohan J. Dutta-Bergman. 2004. Primary sources of health information: Comparisons in the domain of health attitudes, health cognitions, and health behaviors. *Health Commun.* 16, 6 (2004), 273–288.
[16] Lauren Eskreis-Winkler, Ayelet Fishbach, and Angela L. Duckworth. 2018. Dear abby: Should I give advice or receive it? *Psychol. Sci.* 29, 11 (2018), 1797–1806. https://doi.org/10.1177/0956797618795472
[17] Lauren Eskreis-Winkler, Katherine L. Milkman, Dena M. Gromet, and Angela L. Duckworth. 2019. A large-scale field experiment shows giving advice improves academic outcomes for the advisor. *Proc. Natl. Acad. Sci. U.S.A.* 116, 30 (2019), 14808–14810. https://doi.org/10.1073/pnas.1908779116
[18] Amanda L. Graham, George D. Papandonatos, and Kang Zhao. 2017. The failure to increase social support: It just might be time to stop intervening (and start rigorously observing). *Transl. Behav. Med.* 7, 4 (2017), 816–820.
[19] Nir Grinberg, P. Alex Dow, Lada A. Adamic, and Mor Naaman. 2016. Changes in engagement before and after posting to Facebook. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 564–574.
[20] Joe Hair, Carole L. Hollingsworth, Adriane B. Randolph, and Alain Yee Loong Chong. 2017. An updated and expanded assessment of PLS-SEM in information systems research. *Industr. Manage. Data Syst.* (2017), 442–458. https://doi.org/10.1108/IMDS-04-2016-0130
[21] Joseph F. Hair Jr, G. Tomas M. Hult, Christian Ringle, and Marko Sarstedt. 2016. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Sage.
[22] Jamie Hartmann-Boyce, Lindsay F. Stead, Kate Cahill, and Tim Lancaster. 2014. Efficacy of interventions to combat tobacco addiction: Cochrane update of 2013 reviews. *Addiction* 109, 9 (2014), 1414–1425.
[23] Jay Heizer. 2020. *Operations Management: Sustainability and Supply Chain Management*. Pearson Canada, Ontario.
[24] Thaddeus A. Herzog and Christopher O. Blagg. 2007. Are most precontemplators contemplating smoking cessation? Assessing the validity of the stages of change. *Health Psychol.* 26, 2 (2007), 222.
[25] Adrian Holzer, Bruno Kocher, Samuel Bendahan, Isabelle Vonèche Cardia, Jorge Mazuze, and Denis Gillet. 2020. Gamifying knowledge sharing in humanitarian organisations: A design science journey. *Eur. J. Inf. Syst.* 29, 2 (2020), 153–171.
[26] Burcu Kayhan Tetik, Işılay Gedik Tekinemre, and Servet Taş. 2020. The effect of the COVID-19 pandemic on smoking cessation success. *J. Commun. Health* 46, 2 (2020), 471–475. https://doi.org/10.1007/s10900-020-00880-2
[27] Vinoth Kumar Kalidoss and Satvinder Singh Bakshi. 2020. COVID-19 and smoking: An association requiring urgent attention. *J. Smok. Cess.* 15, 4 (2020), 223–223. https://doi.org/10.1017/jsc.2020.26
[28] Samuli Laato, A. K. M. Najmul Islam, Muhammad Nazrul Islam, and Eoin Whelan. 2020. What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? *Eur. J. Inf. Syst.* 29, 3 (2020), 288–305.
[29] Chei Sian Lee, Long Ma, and Dion Hoe Lian Goh. 2011. Why do people share news in social media? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 129–140. https://doi.org/10.1007/978-3-642-23620-4_17
[30] Menghang Li, Nischal Kakani, Chuqin Li, and Albert Park. 2020. Understanding cannabis information on social media: Examining tweets from verified, regular, and suspended users. In *IEEE International Conference on Healthcare Informatics*. 1–10.
[31] Yuting Lin, Carina Tudor-Sfetea, Sarim Siddiqui, Yusuf Sherwani, Maroof Ahmed, and Andreas B. Eisingerich. 2018. Effective behavioral changes through a digital mHealth app: Exploring the impact of hedonic well-being, psychological empowerment and inspiration. *J. Med. Internet Res.* 6, 6 (2018), e10024.
[32] Bernd Löwe, Kurt Kroenke, and Kerstin Gräfe. 2005. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J. Psychosom. Res.* 43, 4 (2005), 163–171. https://doi.org/10.1016/j.jpsychores.2004.09.006
[33] B. Mabey. 2015. pyLDAvis. Retrieved from https://pyldavis.readthedocs.io/en/latest/.
[34] Carol A. Maher, Lucy K. Lewis, Katia Ferrar, Simon Marshall, Ilse De Bourdeaudhuij, and Corneel Vandelanotte. 2014. Are health behavior change interventions that use online social networks effective? A systematic review. *J. Med. Internet Res.* 16, 2 (2014), e2952.
[35] Theresa M. Marteau and Hilary Bekker. 1992. The development of a six–item short–form of the state scale of the Spielberger State–Trait Anxiety Inventory (STAI). *Br. J. Clin. Psychol.* 44, 3 (1992), 301–306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x
[36] Carrie Moore and Lisa Chuang. 2017. Redditors revealed: Motivational factors of the Reddit community. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.
[37] Ghee-Young Noh, Sun Young Lee, and Jounghwa Choi. 2016. Exploring factors influencing smokers' information seeking for smoking cessation. *J. Health Commun.* 21, 8 (2016), 845–854.

[38] Jeni Paay, Jesper Kjeldskov, Umachanger Brinthaparan, Lars Lichon, Stephan Rasmussen, Nirojan Srikandaraja, Wally Smith, Greg Wadley, and Bernd Ploderer. 2014. Quitty: Using technology to persuade smokers to quit. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational.* 551–560.

[39] Zizi Papacharissi and Alan M. Rubin. 2000. Predictors of Internet use. *J. Broadcast. Electr. Media* 44, 2 (2000), 175–196.

[40] Albert Park and Mike Conway. [n.d.]. Tracking Health Related Discussions on Reddit for Public Health Applications. Retrieved from http://www.reddit.com.

[41] Albert Park, Andrea L. Hartzler, Jina Huh, Gary Hsieh, David W. McDonald, and Wanda Pratt. 2016. "How did we get here?": Topic drift in online health discussions. *J. Med. Internet Res.* 18, 11 (2016), e6297.

[42] Namsu Park, Kerk F. Kee, and Sebastián Valenzuela. 2009. Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *Cyberpsychol. Behav.* 12, 6 (2009), 729–733. https://doi.org/10.1089/cpb.2009.0003

[43] Pooja Patwardhan. 2020. COVID-19: Risk of increase in smoking rates among England's 6 million smokers and relapse among England's 11 million ex-smokers. *BJGP Open* 4, 2 (2020).

[44] Michael J. Paul and Mark Dredze. 2011. You are what you Tweet: Analyzing Twitter for public health. In *International AAAI Conference on Weblogs and Social Media (ICWSM'11)*, 265–272.

[45] James O. Prochaska, Carlo C. DiClemente, and John C. Norcross. 1993. In search of how people change: Applications to addictive behaviors. *Addict. Nurs. Netw.* (1993), 2–16.

[46] James O. Prochaska, Colleen A. Redding, Kerry E. Evers, et al. 2015. The transtheoretical model and stages of change. *Health Behav.: Theory Res. Pract.* (2015), 125–148.

[47] James O. Prochaska, Colleen A. Redding, Kerry E. Evers, et al. 2015. The transtheoretical model and stages of change. *Health Behav.: Theory Res. Pract.* (2015), 97–122.

[48] Sheizaf Rafaeli, Tsahi Hayat, and Yaron Ariel. 2008. Virtual knowledge-building communities. In *Encyclopedia of Networked and Virtual Organizations.* IGI Global, 1792–1796.

[49] Nicola J. Roberts, Susan M. Kerr, and Sheree M. S. Smith. 2013. Behavioral interventions associated with smoking cessation in the treatment of tobacco use. *Health Serv. Insights* 6 (2013), 79–85.

[50] Rajani S. Sadasivam, Rebecca L. Kinney, Stephenie C. Lemon, Stephanie L. Shimada, Jeroan J. Allison, and Thomas K. Houston. 2013. Internet health information seeking is a team sport: Analysis of the Pew Internet Survey. *Int. J. Med. Inf.* 82, 3 (2013), 193–200.

[51] Lamyae Sardi, Ali Idri, and José Luis Fernández-Alemán. 2017. A systematic review of gamification in e-Health. *J. Biomed. Inf.* 71 (2017), 31–48.

[52] Marko Sarstedt, Christian M. Ringle, and Joseph F. Hair. 2017. Partial least squares structural equation modeling. In *Handbook of Market Research*, 1–40.

[53] Ralf Schwarzer and Aleksandra Luszczynska. 2008. How to overcome health-compromising behaviors: The health action process approach. *Eur. Psychol.* 16, 3 (2008), 141–151. https://doi.org/10.1027/1016-9040.13.2.141

[54] Saul Shiffman, Sarah E. Brockwell, Janine L. Pillitteri, and Joseph G. Gitchell. 2008. Use of smoking-cessation treatments in the United States. *Am. J. Prevent. Med.* 34, 2 (2008), 102–111.

[55] Charles D. Spielberger, R. L. Gorsuch, R. Lushene, P. R. Vagg, and G. A. Jacobs. 1970. State-trait anxiety inventory for adults: Self-evaluation questionnaire. (1970), 0–75.

[56] L. F. Stead, A. J. Carroll, and T. Lancaster. 2017. Group behaviour therapy programmes for smoking cessation (Review). *Cochr. Datab. Syst. Rev. Group* (2017).

[57] Acar Tamersoy, Duen Horng Chau, and Munmun De Choudhury. 2017. Analysis of smoking and drinking relapse in an Online community. In *Proceedings of the International Conference on Digital Health.* 33–42.

[58] Marjolein Verbiest, Evelyn Brakema, Rianne Van Der Kleij, Kate Sheals, Georgia Allistone, Siân Williams, Andy McEwen, and Niels Chavannes. 2017. National guidelines for smoking cessation in primary care: A literature review and evidence analysis. *NPJ Prim. Care Respir. Med.* 27, 1 (2017), 1–11.

[59] Sheryl L. Warttig, Mark J. Forshaw, Jane South, and Alan K. White. 2013. New, normative, English-sample data for the short form perceived stress scale (PSS-4). *J. Health Psychol.* 18, 12 (2013), 1617–1628.

[60] Robert West. 2005. Time for a change: Putting the Transtheoretical (Stages of Change) Model to rest. *Addiction* (2005), 1036–1039.

[61] Robert West, Martin Raw, Ann McNeill, Lindsay Stead, Paul Aveyard, John Bitton, John Stapleton, Hayden McRobbie, Subhash Pokhrel, Adam Lester-George, and Ron Borland. 2015. Health-care interventions to promote and assist tobacco cessation: A review of efficacy, effectiveness and affordability for use in national guideline development. *Addiction* 110, 9 (2015), 1388–1403.

[62] J. Lee Westmaas, Jeuneviette Bontemps-Jones, and Joseph E. Bauer. 2010. Social support in smoking cessation: Reconciling theory and evidence. *Nicotine Tobacco Res.* 12, 7 (2010), 695–707.

[63] WHO. 2013. Global action plan for the prevention and control of noncommunicable diseases 2013–2020. *Bull. World Health Org.* (2013).

[64] WHO. 2020. Tobacco. Retrieved from https://www.who.int/news-room/fact-sheets/detail/tobacco.
[65] Ken Kwong-Kay Wong. 2013. Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS. *Market. Bull.* 24, 1 (2013), 1–32.
[66] S. U. E. Ziebland and Sally Wyke. 2012. Health and illness in a connected world: How might sharing experiences on the internet affect people's health? *Milbank Quart.* 90, 2 (2012), 219–249.

Received 27 January 2021; revised 28 June 2022; accepted 15 September 2022