MASTER THESIS

Understanding the message functions in health communication, promotion and pubic engagement on Twitter: An exploratory analysis of the SunSmart campaign

Ying-Ling Pan

Twente University

Study: Master Business Administration (Double Degree)

Track: Entrepreneurship & Innovation, Strategy

Student No: s2004224 / 386208

E-Mail: y.l.pan@student.utwente.nl

Date: 31. 08. 2018

University of Twente Supervisor   **TU Berlin Supervisor**

Anna Priante (MSc.)   Dr. Natalia Strobel

Dr. Michel L. Ehrenhard

UNIVERSITY OF TWENTE.

Acknowledgements

During this research, I have received support and guidance from many people. Therefore I would like to take the time to express my sincere gratitude and immense appreciation to at least some of them.

First of all, I would like to thank my supervisors, Anna Priante (M.Sc.) and Dr. Michel L. Ehrenhard. Without their knowledge and guidance, this work would not have been possible. Ms. Priante (M.Sc.), as my first supervisor, always provides me with prompt and constructive feedback which has been essential to any progress that I make. She offers clear guidance and suggestions whenever necessary and her advice inspires me and has given me confidence in this research as well as academic guidance for pursuing excellence in my work. My special thanks go to Dr. Michel L. Ehrenhard who has shown the pathway for this research in the first place and has provided me with possible directions to determine my work.

Last but not the last, I offer thanks to my family and friends for their continued support and encouragement throughout the process of writing this thesis.

Best Regards,

Ying-Ling Pan

August 2018

Abstract

Background. As the mortality of skin cancer has risen rapidly over the recent decades, skin health organisations largely use social media as a communication tool to promote health campaigns and encourage participation. However, little is known about the specific approach to foster engagement via tweets as a form of health communication and promote health campaigns to engage the public. By focusing on the SunSmart skin health campaign on Twitter, this study aims to investigate how the communication during the campaign is characterised in terms of the functions of messages, to what extent the use of these messages can create public engagement, and how message contents play out among the functions.

Methodology. By focusing on the SunSmart health campaign on Twitter, this study adopts a multi-method approach. First, a descriptive statistical analysis is used to understand whether levels of engagement among types of users and message functions differ. Second, Natural Language Processing (NLP) is adopted for developing a codebook in which four message functions manifested from the SunSmart data are identified. Third, content analysis is used to manually classify each tweet to different user types and message functions. Last, by using Natural Language Processing (NLP) and the hashtag visualisation thematic analysis, we further explore whether the composition of content (i.e., keywords & thematic topics) among message functions differ.

Results. Using the 2014 SunSmart health campaign on Twitter as an empirical context and on the basis of comparison between individuals and organisations (i.e., the public), results show that individual users are more engaged in the SunSmart campaign on Twitter than organisations did. In addition, we find the levels of engagement among the four main message functions between individuals and organisations differ. At the content level, results show that utilisation of keywords and thematic topics among different message functions generally differ among individuals and organisations.

Contributions. This study offers contributions to research on media studies, health communication, and health campaign marketing. Practically, the results provides with insight on strategic health communication and marketing campaigns.

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

1. INTRODUCTION

1.1 Situation & Problem Statement

Skin cancer has risen rapidly and become the most common cancer in Australia. According to the Australia Cancer Council report, there were 13,134 Australians diagnosed with melanoma in 2014 and 2,162 people died from skin cancer in 2015. Around two thirds of Australians will be diagnosed with skin cancer by the age of 70, and the problem is getting severe as well in other areas. Skin cancer occurs mainly because of overexposure to ultraviolet (UV) radiation from the sun (SunSmart, 2018). The positive side is that 95% of skin cancers can be treated after early detection. As a result, public health organisations are managing to raise awareness of the dangers of sun exposure and promote skin protection activities via health campaigns.

To reach out to a broad public audience, social media serves as a cost-effective, efficient, powerful health communication and promotion tool for organisations as well as creating opportunities for individuals to disseminate health messages (Bail, 2016; Heldman, Schindelar, & Weaver, 2013; Moorhead et al, 2013; Park, Reber, & Chon, 2016; Park, Rodgers, & Stemmle, 2013). That also makes social media an informative venue that offers researchers critical, insightful perspectives on a wide range of issues including health (Bail, 2016; Heldman et al., 2013; Steinert-Threlkeld, 2017). However, using a social media platform to deliver a message is not a problem, rather how to engage a target audience is (Schultz & Peltier, 2013). Health organisations mostly focus on one-way communication (Bortree & Seltzer, 2009; Rybalko & Seltzer, 2010; Waters & Jamal, 2011; Xifra & Grau, 2010), while an understanding of two-way communication is critical to explore audiences' tastes (Heldman et al., 2013; Neiger, Thackeray, Burton, Thackeray, & Reese, 2013a; Thackeray, Neiger, Burton, & Thackeray, 2013a).

Social media enable public to engage in messages through different engagement tools (Cho, Schweickart, & Haase, 2014), such as *Like, Retweet,* and *Comment* on Twitter, which can be viewed as the two-way communication activities on social media. As the pervasiveness of social media grows, the definition of engagement on social media can be seen as the continuous interaction. Therefore, an investigation of two-way communication activities can offer insights into the public's levels of engagement.

The engagement needs to be embraced for a couple of reasons. First of all, it has the potential impact on health behavior change (Healdman et al., 2013). The engagement on social media reveals the people give weight to messages on social media, and then act of response,

which this paper suggests that is an antecedent of attitude and behavior to participate in a health program physically.

Additionally, the findings of prior studies support the fact that social media plays an essential role in marketing channels and can be a useful marketing tool to influence people's behavior (Bruhn, Schoenmueller, & Schäfer, 2012; Heldman et al., 2013). In the context of health promotion, the engagement on social media has been characterised as connections between people that contribute to a common good (Neiger, Thackeray, Burton, Giraud-Carrier, & Fagen, 2013b). An understanding of the benefits of engagement to health campaign promotion allows a health organisation to identify the health information needs of users (Heldman et al., 2013; Neiger et al., 2013a) and further helps health organisations to craft a marketing plan that tailors compelling messages (Leek, Houghton, & Canning, 2017). Marketing strategies on social media may facilitate people's online engagement, enable people to acquire more knowledge, information about health and further advance their health promotion (Thackeray, Neiger, & Keller, 2012).

However, while a range of papers have suggested that using social media platforms to support engagement (Ashley & Tuten, 2015; Neiger et al, 2013; Swani, Brown, & Milne, 2014; Swani, Milne, Brown, Assaf & Donthu, 2017), little is known about how to effectively use social media to communicate with the public (Burton & Soboleva, 2011; Lacoste, 2016; Neiger et al. 2013a; Park et al., 2016; Siamagka, Christodoulides, Michaelidou, & Valvi, 2015; Wiersema, 2013); a thorough understanding of how to effectively utilise social media requires further research (Lovejoy, Waters, & Saxton, 2012).

In regard to health communication on social media, previous studies show that identifying different categories of messages from the content can favour us in-depth insights (Burton, et al., 2013; Chew and Eysenbach, 2010; Hambrick et al., 2010; Harris, Mueller, Snider, & Haire-Joshu, 2013; Leek et al., 2017; Lovejoy et al., 2012; Neiger et al., 2013a ; Park et al., 2016; Sullivan et al, 2011; Thackeray et al., 2013a; Van den Heerik, van Hooijdonk, Burgers, & Steen, 2017). For delivering an effective message, the purpose of a message needs to address audiences precisely. Researchers have pointed out the importance of message functions (i.e., the purpose of messages) manifested from messages and content of messaging (Lovejoy et al., 2012; Leek et al., 2017; Park et al., 2016; Sullivan et al, 2011), whereas analysis of the specific content of communicated messages is rare (Waters & Jamal, 2011). Research literature show not much information about how to specifically frame a tweet to fulfil the

message function to increase public engagement (Lovejoy et al., 2012; Neiger et al., 2013a; Park et al., 2016).

As a result, this paper aims at understanding how the communication during a skin health campaign is characterised by identifying the functions of messages on social media-Twitter, whether different message functions can create different public online engagement, and specifically how they differ in contents.

Therefore, on the basis of comparison between individuals and organisations, this study examines the level of engagement among these two types of users and message functions. In this way, an understanding can be gained of which types of users are more engaged in the SunSmart campaign and which message functions are more effective. Thereafter, this study further explores content variation of message function to gain an in-depth insight into the framing of a particular type of message to foster public engagement.

1.2 Research Goals & Research Questions

1.2.1 Research Goals

By scrutinizing tweets, this study aims to explore how the communication during the campaign is characterised by identifying the types of messages, examine whether the public's (i.e., organisations and individuals) levels of engagement (i.e., retweets) in the SunSmart campaign differ by message functions. Furthermore, what composition (i.e. keywords and thematic topics) constitute different functions, is investigated via the following stages:

(1) Identify whether the levels of engagement (i.e. number of retweets) with the SunSmart campaign differ per different senders (i.e. individuals and organisations)

(2) Examine whether levels of engagement (i.e. number of retweets) in the SunSmart campaign differ per message functions for individuals and organisations.

(3) Examine whether the specific language used, looking specifically at keywords and thematic topics, manifest in varied message functions for individuals and organisations.

1.2.2 Research Questions

To fulfil our research goals, this study poses the following research questions:

First of all, since individuals and organisations are two distinctive types of senders, distinguishing the types of these two users can offer an overall comparative insight in regard to what extent that these two parties use Twitter to post messages and their levels of engagement

in SunSmart. Consequently, as a first step to investigate the level of engagement (i.e. retweet) of these two types of users, the first research question is:

RQ1: How do the levels of engagement differ by individuals and organisations during the SunSmart campaign?

Moreover, as content posted on Twitter might have different functions, such as *information sharing, problem solving,* and *public relations* (Leek et al., 2017), some tweets may generate more engagement than others. Therefore, after identifying the message functions according to the model of Leek et al. (2017), investigate how message functions generate public engagement. Hence, the second question is:

RQ2: To what extent do message functions observed from the manifest content of tweets in connection with the SunSmart campaign differ regarding the level of engagement among individuals and organisations?

Furthermore, to understand how to construct effective message categories, the composition of message functions in regard to semantic content (i.e., word choice) and theme can assist in disseminating more specific information. In this study, two items are subject for our content analyses: keywords and thematic topics. The keywords analysis is conducted to point out the most significant words in a particular message function category and a thematic topics analysis can reveal primary topic groups and their patterns in message functions. Hence, the further research questions RQ3a and RQ3b are raised:

RQ3a: How do the keywords of tweets regarding the SunSmart campaign differ in regard to message functions among individuals and organisations?

RQ3b: How do the thematic topics of tweets related to the SunSmart campaign differ in regard to message functions among individuals and organisations?

1.3 Research Motivation

Social media as a communication tool in literature show little information about two-way communication for engagement

To raise the public's awareness of skin health issues and to reduce the number of deaths from skin cancer, health organisations have begun using social media for health information communication and campaign promotion (Amina Jama Mahmud et al., 2013; Bruhn et al., 2012;

Jha, Lin, & Savoia, 2016; Leek , et al., 2017; Lovejoy et al, 2012; Neiger et al., 2013a; Park, et al., 2016; Priante, Need, Van den Broek, & Hiemstra, 2018a; Rus & Cameron, 2016; Smaldone, et al., 2015; Wu, et al., 2016). While research works suggest that using social media to increase public engagement (Ashley & Tuten, 2015; Neiger et al, 2013a; Swani et al., 2014; Swan et al., 2017), a thorough understanding of effectively utilising social media requires further research (Lovejoy, Waters & Saxton, 2012). It is found that not much literature focus on the two-way communication in which we can have more in-depth insights into the public's engagement and campaign marketing (Heldman et al., 2013; Neiger et al., 2013a).

Social media as the health communication in prior works show the message function and content is the key, but a deeper understanding of patterns among message functions lack

The research literature has shed light on the importance of content analysis and message function (i.e., purpose of tweets) (Burton, et al., 2013; Leek, et al., 2016; Lovejoy, et al., 2012; Naaman, Boase, & Lai, 2010; Park et al, 2016; Sriram, Fuhry, Demir, Ferhatosmanoglu, & Demirbas, 2010; Sullivan et al., 2011) ; however, not many researchers identify the patterns and structures of message functions; that is, there is little information about how to specifically frame a tweet to fulfil a certain type of message function. Hence, an understanding of patterns and structures of message functions and levels of engagement can provide significant insights into how to create the most intriguing content for audiences on social media.

A new analysis perspective considers the general user: previous papers with results focused on engaging active users, which contributed less to expanding their general audience base

Most research on message functions and public online engagement in the health sector focuses on the more active Twitter users who follow a healthcare organisation's Twitter account as a target population (Leek et al., 2017; Neiger et al., 2013a; Park et al., 2016; Thackeray et al., 2013a). However, active participators such as SunSmart's followers, usually already had relatively high interest in the SunSmart campaign more so than other potential audiences, and may have different preference for health issues. Thus, this study focuses on all users by using a complete SunSmart data that can contribute insights from a different perspective.

1.4 Outline of this Thesis

This master thesis is divided into *six chapters* which are structured as follows to answer the research questions:

     Chapter 1 explains the importance of conducting this research by underlining the problem of low-levels of public awareness in regard to the issues of skin health. *Chapter 2* is a relevant research literature review that includes an explanation of social media (i.e. Twitter) and its usage by health organisations, followed by a definition of public online engagement, connecting engagement to how health organisations use social media for marketing a campaign as well as relevant points from previous studies in regard to message functions in health communication. Thereafter, *Chapter 3* illustrates the research design comprising the conceptual model, data scope, and a combined approach by using the descriptive statistical analysis, qualitative content analysis as well as computational text analysis tools that have been conducted as an examination of our findings. *Chapter 4* analyses the research results and highlights some main points. The results indicate whether the levels of engagement among individuals and organisations differ, whether message function varieties lead to their degree of public engagement and whether keywords and thematic topics differ among message categories. *Chapter 5*, highlights key findings and, discusses related future research possibilities as well as the limitations of this study. *Chapter 6*, the conclusions summarises how to strategically use tweets as well as frame effective messages to foster public engagement.

2. LITERATURE REVIEW

2.1 Use of Twitter for communication

2.1.1 Social Media Platform: Twitter

Social media is the collective of online communication channels that facilitate exchange of ideas, information-sharing, and interaction via virtual communities (Heldman et al., 2013). While there is not a consistent definition due to its variety, it is commonly described as interactive internet-based applications with user-generated content, such as text posts, comments or other data generated through all online interactions (Fox, 2011; Osborne-Gowey, 2014).

According to the Pew Research Centre 2018 survey, social media use is widespread among internet users (Smith & Anderson, 2018). In many of the advanced economies surveyed, at least two-thirds of all adults in the U.S., Australia, South Korea, Canada, Israel and Sweden use social media. These high usage rates of social media are also found in emerging and developing economies.[1] Statia statistics shows that there are 2.4 billion social media users all around the world, and Twitter users account for at least 330 million (Statia, 2017).

Today, the social media platform has become a powerful communication tool (Bail, 2016; Heldman et al., 2013; Park et al., 2013; Park et al, 2016; Moorhead et al., 2013). People share their lives via many platforms like Facebook, Twitter, and Instagram, which makes the influence of social media increasingly great. Around 71 % of internet users are social network users and these figures are expected to grow (Statia, 2017). Across 39 countries of the advanced economies, on an average of 53% citizens say they use online social networking sites like Facebook or Twitter.[2] Social media platform, such as Twitter, serve as a tool for organisations to promote campaigns, disseminate campaign-related information, and motive people to discuss, share their opinions and participate in activities. (Bravo & Hoffman-Goetz, 2017; Jacobson & Mascaro, 2016; Luo & Smith, 2015).

In addition, social media offers a platform for bottom-up discovery users' opinions. The contents of social media from users also offer traces regarding how to tailor conversations that

---

[1] http://www.pewglobal.org/2018/06/19/social-media-use-continues-to-rise-in-developing-countries-but-plateaus-across-developed-ones/#table

[2] http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/

can fit target audiences' needs and knowledge levels (Neiger et al., 2013a) Twitter provides researchers with a large database (Steinert-Threlkeld, 2017), that also makes social media an informative venue that offers researchers critical, insightful perspectives on a wide range of issues including health (Heldman et al, 2013). Particularly, social networking platforms offer researchers a means to understand the structure and patterns of conversations (Steinert-Threlkeld, 2017). Therefore, the value of social media is further enhanced when organisations utilise these platforms to create ongoing conversations and dialogue with their audiences (Heldman et al., 2013; Priante et al, 2018a ; Thackeray et al., 2012).

A microblog, among different types of social media, is defined as websites that are particularly useful for sharing time-sensitive information and opinions by using less than 200 characters (Gallaugher, 2013). Because a typical human being's attention span is limited, a page filled with massive blocks of text probably is a daunting prospect for many audiences and likely limits their ability to pay attention such verbosity. From several studies, Twitter, the one of the most widely-used microblog which limits each "Tweet" to 140 characters, has been recognised as an effective communication and engagement tool more so than other social media by yielding several benefits. Firstly, a vast majority of Twitter's data are publicly accessible again more so than other platforms such as Facebook (Gallaugher, 2013). Secondly, the platform is research friendly. Through its Application Programming Interference (API), researchers are allowed to import vast amounts of data rapidly. With a large, easy and freely accessible dataset, the latest statistic shows around 500 million tweets are being posted daily (Steinert-Threlkeld, 2017). In addition, with user-generated content, Twitter provides real-time information and two-way communication that can facilitate organisational communication by building a dialogic loop (Heldman et al, 2013). It provides users with an interactive platform through the frequent posting of short messages and a number of interactive functions (Heldman et al, 2013). McCormick et al., (2013) assert that the real-time setting of Twitter enables researchers to observe human behaviour without influencing the behaviour of interest. Also, Twitter is extremely cost-effective and makes scalability possible (McCormick et al., 2013).

Furthermore, Twitter is an ideal platform to target young audiences. Skin cancer can be treated after early detection so that the earlier the public can be aware and participate in skin health prevention activities, the higher the possibility that they can avoid this disease. According to a survey of social media use from Pew Research Center 2018, there are substantial differences in social media use by different age demographic groups, and Twitter users are

relatively younger.[3] This result is also supported by the digital marketing agency, Ominicore, who published data indicating that around 37% of Twitter users are between the ages of 18 and 29, while 25% of users are 30-49 years old.[4] Men and women use Twitter in almost equal numbers (Greenwood et al., 2016). These figures reflect the fact that Twitter is suitable to be a research venue with a young, gender-balanced sample population (Steinert-Threlkeld, 2017).

2.1.2 Literature show little information about Two-way Communication

Using a social media platform to deliver a message is not a problem, rather how to engage a target audience is

As social media has become widely-used, it has been seen as an influential communication tool for the health sector in health communication and campaign promotion. Healthcare organisations have recognised the value of, and embraced the use of, social media for disseminating information (Heldman et al., 2013; Neiger et al., 2013a). However, in this age of information overload, which is in part created by social media, messaging in regard to health concerns may be easily ignored or unable to draw the attention of target audiences. For many health organisations or campaigns, delivering messages via social media is not a problem, but how to engage the public via post content is (Schultz & Peltier, 2013).

In regard to how to best to use Twitter for health communication, there are different opinions. Health organisations often adopt one-way communication via social media by simply disseminating their message rather than engaging audiences via a two-way communicative approach (Bortree & Seltzer, 2009; Lovejoy & Saxton, 2012; Rybalko & Seltzer, 2010; Waters & Jamal, 2011; Xifra & Grau, 2010). Some papers suggest that one-way communication adopted by traditional mass media campaigns on Twitter is still the prevalent communication approach (Waters & Jamal, 2011; Xifra & Grau, 2010). Researchers state that it provides a point of contact to attract potential customers' attention and is essential to initiate a dialog (Lovejoy & Saxton, 2012; Waters & Jamal, 2011). Some studies point out that one-way communication is particularly important in some circumstances, such as disease epidemics or disaster emergencies (Hughes & Palen, 2009; Merchant, Elmer & Lurie, 2011; Smith, 2010).

---

[3] http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/
[4] https://www.omnicoreagency.com/twitter-statistics/

Two-way communication on social media involves more engagement

However, little evidence shows that using Twitter as a one-way communication tool for health information dissemination is effective in improving health status (Neiger et al., 2013a). Instead, one-way communication eliminates opportunities to gain a better understanding of audiences. It can be argued that engagement is the key to successful social media efforts (Korda & Itani, 2013; Li & Bernoff, 2011; Safko, 2010). Without continuous interaction with target audience, it is hard to foster conversations and may hinder publics' engagement (Scott, 2015). Organisations also risk losing its audiences (Heldman et al., 2013). Furthermore, continuous interactions with audiences probably encourage actions on health campaigns participation (Neiger et al., 2013a). Since social media offer a range of communication tools to engage publics, such as *Like, Share* and *Comment* on Facebook (Cho et al., 2014), the two-way communication activities involve more engagement can favour us the in-depth insights into the public's engagement.

2.2 Public online engagement of a health campaign

2.2.1 Social media broadens interactive communication for public engagement

The importance of public engagement on social media

Social media engagement needs to be embraced for a couple of reasons. First, it has the potential impact on health behaviour change (Healdman et al., 2013). The engagement on social media reveals that people give weight to the message with a cognitive, affective commitment to the tweet message on Twitter, then act of sharing information, which this paper suggests that is an antecedent of attitude and behaviour to participate in a health program physically. The engagement needs to be further studied to know how to foster engagement via social media channels to encourage health behaviour (Healdman et al., 2013).

Moreover, an understanding of engagement allows health organisation to identify the health information needs of users (Heldman, 2013; Neiger et al., 2013a), to craft a marketing plan that composes intriguing messages to increase public's levels of engagement (Leek et al., 2017) and advance the public's health.

Definition of engagement

Although the term "engagement" has been widely used in literature, a benchmark definition is still lacking. Research has shown that engagement can be explained in various forms. From the

perspective of marketing, engagement can be viewed as repeated interactions that strengthen a consumer's emotional, psychological, or physical investment in a brand (Sedley, 2008). Some researchers argue that it is not limited to transitional behaviour and it can be seen as measurements of information sharing in the decision-making process during value co-creation. Co-creative experience with a focal agent can be built up and exist as a dynamic and iterative process through interaction (Brodie, Hollebeek, Jurić, & Ilić, 2011). Brodie et al., (2011) explain engagement as the participant's specific interaction in virtual communities (Brodie et al., 2011). Van Doorn, et al., (2010) describe engagement in terms of a psychological state from interactive (Van Doorn et al, 2010). While the interpretations of engagement are varied, engagement generally refers to audience interactions and their experience with stakeholders.

As the pervasiveness of social media grows, non-profit organisations' ability to communicate with stakeholders such as volunteers and the public is significantly enhanced (Heldman et al., 2013). Online interactions have become multifaceted and critical dimension to organisations' performance (Lovejoy & Saxton, 2012). The online conversations on social media highly affect customers' purchase intentions (Tsimonis & Dimitriadis, 2014). The improvement of digital communication has extended the scope of information exchange, which makes interpersonal contact and information sharing between relevant parties easier (Tuten & Ashley, 2016; Swani et al., 2017). Social media enlarges the territory of interactive communication from the traditional definition of engagement to that of the digital engagement definition. Saxton & Waters (2014) explain online engagement as the public's involvement in organisational activities by responding via social media (Heldman, 2013; Saxton & Waters, 2014). Therefore, in the setting of social media, engagement can be seen as public participation in continued interaction via a social media platform. That is, engagement can be manifested via actions such as message sharing.

2.2.2 Twitter: venue for interactive communication & investigating public engagement

Kent and Taylor (2002) identify online communication in particular as an ideal avenue for fostering dialogue (Kent & Taylor, 2002). The findings of other studies show how organisations use online communication to facilitate interactive communication with the public via social platforms (Bortree & Seltzer, 2009; Burton & Soboleva, 2011; Hackler & Saxton, 2007). Twitter's potentially contingent interactive messages can assist organisations in communicating with other users. Typically, the communication tools on Twitter include:

(1) Mention: The Mention function can create interactivity by directing a message to a user's account with the symbol "@" or "at" in English. Thus, posting a tweet with the mention symbol "@" before a username directs that message to a specific user. Organisations can use the function to communicate with other users by including the "@" symbol and a Twitter username in a message. When using the Mention function in a public message, a dialogue is established between an organisation and the user, and the message is also visible to their followers. This feature enables organisation to draw audiences' attention and stimulation conversations (Boyd, Golder & Lotan, 2010). For instance, the following tweet comprised a question and mentioned the user @Katieclift. The user "@Katieclift" can see the mention and is able to give a response to the question.

- 'Do you live in a #melanoma hotspot? @Katieclift explains http://t.co/yid2CcUNN5 #sunsmart #dailyshade #skincancer http://t.co/ejmknBT9Tj'

(2) Hashtag: Another critical feature is the hashtag, denoting the pound sign (#) before a relevant keyword or phrase (e.g., #SunSmart). Hashtags can be put in any position for user-defined topics in a tweet. This function can group conversations by topics and allows people to follow topics they are interested in easily. Users can simply click on a hashtagged word to find out what other tweets include the same hashtag. The use of hashtags makes searching for information easier by categorising conversations (Funk, 2011). For example, if a user searches for information about skin, the term "skin" would come with a number of results. However, using hashtag #skin would yield results more relevant to the topic.

Additionally, hashtagged words often become popular or trendy. Health organisations can use hashtags for important topics or include a popular topic to make an even more visible (Funk, 2011). In essence, hashtags are designed to identify the topics of communication. Observing hashtag use among the public we can see what are trends and what topics are emerging as well. A dialog between users and a health organisation can be more easily built up by effective use of hashtags, which enables a topic to travel faster (Heldman et al., 2013).

(3) Retweet: Another communication function is the retweet, a function that allows users to share a tweet, posted by other users, with their own followers. This is a useful function for users to re-post an old tweet to ensure their followers can see it. A retweet can be used to answer public messages to keep a complete dialog or to share a message. When an organisation forwards other users' tweets to share information with followers, it can demonstrate

connections with other organisations or individual users (Lovejoy et al., 2012). The Retweet is a quick way to pass along news and interesting discoveries on Twitter and can be seen as a way by which participants being in a conversation (Boyd et al., 2010). Within the downloaded data, a Retweet gives acknowledgment of the user by adding "RT@[username]" to the beginning of the message to avoid confusion with the Mention function "@". The following example, sent by a football community account AtavusRugby, is displayed to explain how to distinguish a Retweet and a Mention.

- "RT @christombs71: @RyyKayCar @SereviRugby thanks Ryan! #sunsmart #European! It's great here in SoCal though"

This message can be recognised as a Retweet by the "RT @christombs71," which indicates that AtacusRugby shared the user @christombs71's post, and the main text body comprises the "mention" function directing this message to another two users @RyyKayCar and @SereviRugby.

(4) Reply: A reply represents a response to another user's tweet by clicking the "reply" button, which enables the public to participate in dialogic communication (Heldman et al., 2013). If a user replies to another user, a "replying to" mark preceding the tweet is shown on his/her profile page timeline. Clicking on a reply in users' timeline, a full dialogue can be seen.

Among these mechanisms, this study defines engagement as continued interaction with this online community, and the *retweet* served as the indicator. A retweet by others has the advantage of an apparent independent endorsement (Romero et al., 2010; Burton, & Soboleva, 2011). People retweet when they want to spread newly-discovered or strongly resonating information to friends and new audiences and publicly value others' opinions (Boyd et al., 2010). The *retweet* has become the principal mechanism for making information travel on Twitter (Suh, Hong, Pirolli & Chi, 2010), thereby discovering what contents tended to be retweeted by the public can lead to a critical insight into what determines information diffusion.

To sum up, through public messages, an interaction is created between the Twitter account (i.e., an organisation or an individual) and followers and also viewable by anyone following the Twitter account. All these public functions can be employed to commit to creating ongoing communication, which makes Twitter an ideal venue to investigate public engagement (Lovejoy et al., 2012).

2.2.3 Engagement: importance for health organisation to craft marketing strategies

The findings of prior studies support the fact that social media plays an essential role in marketing channels and can be a useful marketing tool to influence people's behaviour (Bruhn et al., 2012; Heldman et al., 2013). Social media can facilitate interaction and communication through sharing information and building dialogic relationships (Heldman et al., 2013; Lovejoy & Saxton, 2012; Luo, & Smith, 2015; Saffer, Sommerfeldt & Taylor, 2013). In the context of health promotion on social media, engagement has been characterised as connections between people that contribute to a mutual benefit (Neiger, Thackeray, Burton, Giraud-Carrier, & Fagen, 2013b).

An understanding of the benefits of engagement to health campaign promotion allows a health organisation to identify the health information needs of users (Heldman et al., 2013; Neiger et al., 2013a). This may be supported by 'Uses & Gratification Theory' (Katz, Blumler, and Gurevitch, 1973/ 1974). This suggests that individuals respond to content that satisfies their needs (Dolan, Conduit, Fahy, & Goodman, 2016; Gao & Feng, 2016). Organisations should identify opportunities to connect with users and create opportunities to engage more audiences, both of which allow them to have opportunities to engage more users (Heldman et al., 2013).

In addition, broadly delivering information without knowing the audience may be inefficient for a health campaign promotion. A successful health campaign relies on the public's participation as well as distributing information. Understanding what content engages the audience in a conversation contributes greater insight into making a successful marketing plan for a health campaign (Heldman et al., 2013). A successful social media campaign enables customers to interact with organisations (Safko, 2010). As a result, having a deeper insight into public engagement is essential for SunSmart marketing strategies.

Because of the important role of social media, social media marketing has been more common in literature (Ashley & Tuten, 2015; Chang et al., 2015; Thackeray et al., 2012;). Nkanunye & Obiechina (2017) state that health promotion provides the capacity to assist people in identifying health needs as well as obtaining resources to achieve change in health (Nkanunye & Obiechina, 2017). By providing thoughts and tools that more effectively reach and impact the target audiences for health campaigns, marketing in the health area can be viewed as increasing the effectiveness of health promotion campaigns (Donovan, 2011). Through marketing activities to increase the level of engagement of the public in health

campaign, people can acquire more knowledge, information, and resources that assist them in identifying health needs and may further improve their health. Marketing strategies by engaging messages on social media may facilitate people's engagement and advance health promotion (Thackeray et al., 2012).

Although many researchers suggest that organisations use social media to increase audience engagement (Ashley & Tuten, 2015; Neiger et al., 2013a ; Swani et al., 2014; Swani et al., 2017), there is little evidence to guide managers in developing a cost-effective strategy for external communication (Burton & Soboleva, 2011; Heldman et al., 2013). To understand whether the levels of engagement by individuals and organisations differ in relation to the SunSmart campaign, RQ1 is asked.

2.3 Previous studies about Message Function in communication for engagement

2.3.1 Type of messages: Message functions serve as a communication ambassador

Owing to the availability and importance of social media, people now have increasingly greater access to health information (Campbell & Craig, 2015). While a range of papers have suggested that using social media platforms to support engagement (Ashley & Tuten, 2015; Neiger et al., 2013a; Swani et al,; Swani et al., 2014; Swani et al., 2017), a thorough understanding of how to effectively utilise social media requires further research (Lovejoy et al., 2012).

To grasp how messages are used for communication on social media, we looked at literature for health communication and found a range of classifications for messages as below. Chew and Eysenbach (2010) investigate how the behaviour of tweeting changed during the H1N1 pandemic crisis. For the contents of tweets, six content categories were identified from the data: *Resources, Direct or indirect personal experiences, Personal reactions, Opinions/ jokes or parodies, Marketing for H1N1-related products, and Unrelated posts.*

Hambrick, Simmons, Greenhalgh & Greenwell (2010) examine the use of Twitter by professional athletes for communicating with fans and other players. The study identifies six categories from the contents of tweets: (1) *Interactivity* (direct communication with fellow athletes and fans), (2) *Diversion* (non-sports-related information), (3) *Information sharing* (insight into athlete's teammates, team or sports etc.), (4) *Content* (including links to pictures, videos or other Web sites), (5) *Fanship* (sports discussion related to teams) and (6) *Promotional* (upcoming games, discounted tickets, giveaways or sponsorships relevant). They

found that "interactivity" is a major category (34%) which may enhance the support of fans (Hambrick et al., 2010).

Harris et al. (2013) aim to understand how health departments use social media to educate and inform the public about diabetes, and three different categories: *Risks, Benefits,* and *Cues to actions* are distinguished.

Thackeray et al. (2013b) study how Twitter is being used during the best-known breast cancer awareness event- Breast Cancer Awareness Month (BCAM). There are eight categories: *Clothing, Fundraiser, Walks, Early detection, Loved ones, Diagnosis, Treatments,* and *Resentment* involving in the tweets identified for the BCAM.

Van den Heerik et al. (2017) investigate how the slogans from the target audience resonate with or deviate from the campaign's original message to get insight in the use of co-creation for slogans in anti-smoking health campaigns. In the study, 11 domains were distinguished from messages on different social networking sites: *Big event, Eating & stimulants, Hobby & Hype, Technology & innovation, Person & group, Sex & relation, School, Transport, Campaign, Personal features, Social norm.*

In addition, Lovejoy et al., (2012) identify three main message functions of tweets: *Information, Community*, and *Action* from among the messages disseminated by the100 largest non-profit organisations (Lovejoy et al., 2012). Among these three functions, *Information* comprises reports, news, facts, and events; *Community* mainly focuses on responses; while *Action* includes tweets that encourage followers to support the organisations' activities. Based on Lovejoy and Saxton's original classifications, Thackeray et al. (2013a) discover what the primary function of Twitter use is among state health departments.

Moreover, Neiger et al. (2013a) retain Lovejoy's and Saxton's definitions of the three categories but replace the "community" by the term "engagement" to examine how local health departments use Twitter to share information, engage with followers, and promote action (Neiger et al., 2013a). The coding was designed to determine the purpose of the tweets. These three categories were also adopted by Park et al. (2016) in which researchers examine how health organisations use social media (i.e., Twitter) in health communication and public engagement.

Leek et al. (2016) published a study regarding the use of Twitter by the health sector in which they also develop three different message functions: *information-sharing*, *public-solving*, and *public relationship* and examine how the tweets of use by health companies differ in the level of behavioural engagement.

Among these literature, types of messages are determined based on the concept of topics or the purpose of messaging. These studies show that identifying different categories of messages from tweets can favour us in-depth insights. For delivering an effective message, the purpose of a message (i.e., message function) needs to address audiences precisely. Organisations can use different message function types to achieve diverse tasks (Leek, Canning, & Houghton, 2016). Message functions can be explained as achieving different communication purposes through delivering different types of messages.

To understand how communication is characterised during a health campaign by the types of messages, what message functions (i.e., purpose of message) can result in higher engagement is what we eager to know; From these message function perspectives on health communication, some categories that are found in both Lovejoy and Leek (Park et al., 2016; Leeks et al., 2017) are presented in this thesis and the details are illustrated in section '3. Methodology' herein somewhat hereafter.

Therefore, to understand whether the levels of engagement differ among different message functions and types of users, more specifically, how effective a message function is to individuals and organisations, the RQ2 is formed.

2.3.3 Message content is critical to fulfil the function of messaging

To fulfil message functions, message content is critical because it very likely influences users' interests and determines if a message may be favoured with a retweet. Understanding the variance of message content in different message functions and levels of engagements also helps social health organisations to frame engaging messages as well as strengthening their brand's social media position (Leek et al., 2017).

Content produced at an increasing rate leads to massive amounts of text in documents that may be analysed (Waldherr, Heyer, Jähnichen, Niekler, & Wiedemann, 2016). Language, serving as a medium that enables psychologists to attempt to understand human beings, can be seen as a reliable way to transform people's thoughts and sentiments into a recognisable form (Tausczik & Pennebaker, 2010). This statement is supported by the research of Leek et al., (2017) who find that diverse content is displayed via different message functions which suggests that the function of a tweet and language should be taken into account when composing a message (Leek et al., 2017).

Language

In addition, the language used in such a short message may play a more important role since linguistic characteristics are more specific to a successful communication to make an action happen (Leek et al., 2017). Although communication tools such as URLs may amplify the interaction between stakeholders on the Twitter infrastructure, e.g., an embedded link can quickly bring the content of a third party's website into the tweeted message; however, message functions require the health organisation to signal via the message text explicitly. Twitter communication mechanisms such as video, photo, and URLs cannot influence public engagement alone (Leek et al., 2017). Therefore, finding out the most relevant word (i.e., keywords) of a type of message function is crucial to fulfil the purpose of the message. Besides, as users can proactively search for information through Twitter's infrastructure, this makes communication via content become particularly important. Thus, to understand how keywords differ among message functions with different levels of engagement for individuals and organisations, RQ3a is asked.

Themes

Furthermore, the topic of a tweet may determine the number of times a message gets retweeted by other users (Boyd et al., 2010). Health organisations engage the public in regard to health-related issues via social media, which connect them with other users with similar interests (Heldman et al., 2013; Lovejoy et al., 2012).

However, although researchers have pointed out the importance of message functions and content to engagement (Heldman et al., 2013; Lovejoy et al., 2012), not much information specifically explaining how to frame tweets using different message functions to increase engagement with the public has been produced. An individual tweet may lack information and not be interesting, but the aggregation reveals interesting patterns about what topics are salient and vary among a given group of people (Steinert-Threlkeld, 2017). Merely identifying a single word or categorising topics among different message functions may not be enough to tell us how to construct the topics. For instance, a topic such as "*Informing about organisational events*" provides a direction but is unable explain how to compose it or indicate a particular event. Therefore, discovering the pattern and structure of topics among a message function can tell us what people are saying and how they construct the topics.

Hashtags, serving as an important feature, convey information that people value. It is worth discovering how they are used in different message functions, observing their patterns

and structure can offer an in-depth insight into how thematic topics form and are varied among message functions (Gründemann, & Burghardt, 2016)

This thesis is mainly inspired by the work of Leek et al., (2017) in which message functions, linguistic characteristics and levels of engagement are discussed. However, how to compose a tweet as described prior, such keywords and thematic topics of message functions, discovered by observing the usage of hashtags among message functions, deserves more attention and discussion due to the insights also previously described. To understand how thematic topics vary among message functions with different levels of engagement for individuals and organisations, the RQ3b is posed.

3. METHDOLOGY

3.1 Research Design

As stated in the prior literature review, the importance of message functions, message content and engagement is considered. This exploratory study is conducted based on a content analysis concept and adopts a multi-method approach combining: a) descriptive statistical analysis b) Natural Language Processing c) content analysis d) hashtags visualisation thematic analysis, which favours an in-depth investigation into content that compels or inspires the public to share via reposting or specifically in this case through a retweet. The details are illustrated in *3.1.1*.

This study's research design is inspired by the paper of Leek, et al., (2017) (referred to hereafter as "Leek" for brevity), which examines how product and service companies use Twitter in the healthcare sector. Appendix A explains Leek's research structure. Leek's study investigates four company's tweet content posted on Twitter that leads to different levels of engagement in terms of likes, retweets and comments. The study firstly determines whether the levels of engagement differ per company types (i.e., service and product) and message functions. Thereafter, a linguistic content analysis was conducted to determine the type of content associated with greater engagement. In Leek's research it is demonstrated that the usage of message functions and content is of importance via linguistic analysis. However, Leek's study is mainly focused on single words within content, yet keywords and thematic topics in a tweet also deserve more attention and can tell us much about the structure of message functions.

3.1.1 Analytic Approach: a multi-method approach

As stated above, to have an in-depth insight, we used a multi-method approach. First, to measure the levels of engagement between individuals and organisations, a descriptive statistical analysis was adopted.

Following, as online content has become more complex due to the interconnection via hyperlinks or hashtags (Hamilton, 2013), merely manual quantitative content analysis for such data is increasingly difficult to carry out due to unstructured data (Waldherr et al., 2016). Computational text analysis has the strength of dealing with a large amount of data, helps to discover patterns in data and deliver rather broader research results (Fass & Turner, 2015). Computational text analysis is a growing area in the social sciences (Grimmer & Stewart, 2013; Lucas et al., 2015). Scholars have begun to explore the greater possibilities offered via computing methods for content analysis, (e.g., Feldman, & Sanger, 2006; He, Zha, & Li, 2013;

Stieglitz & Dang-Xuan. 2013; Stieglitz, Dang-Xuan, Bruns, & Neuberger, 2014). As a result, we used two computational text analysis tools- *Natural Language Processing (NLP)* and *Hashtags visualisation thematic analysis* to scrutinize data and discover patterns from text.

In addition, Stempel (2003) defines content analysis as "a formal system for doing something we all do informally rather frequently- draw conclusions from observations of content" (Stempel, 2003, p.209). The qualitative content analysis that can deliver in-depth research results (Mayring, 2000) is used to annotate the tweets manually in this research.

Therefore, in this study, a multi-method approach that includes *four* different analyses that are utilised in different research "stages," the details are illustrated in 3.1.3 "Revised Model":

(1) Descriptive statistical analysis via SPSS: a comparison of level of engagement (Stage 2 & Stage 3)

(2) Natural Language Processing (NLP) via Python: a multifaceted computational text method that was used to scrutinise data for developing codebook (Stage 3) and identifying keywords (Stage 4).

(3) Content analysis: manually annotation for user types and message functions categorisation (Stage 2 & Stage 3)

(4) Hashtags visualisation thematic analysis via Cortext manager: an automated and computational text analysis tool to find thematic topics by hashtags patterns, structure among message functions. (Stage 4)

3.1.2 Dependent Variable

Twitter defines engagement as the number of times a user interacts with a tweet. For example, the interaction can result from "retweets, favourites, likes, comments, and embedded media" (Twitter). A similar definition also shows up in other health related studies (Leek et al., 2017; Park et al., 2016; Rabarison et al., 2017; Welch, Petkovic, Pardo, Rader, & Tugwell, 2016). Retweeting, is an extremely useful mean for companies to disseminate information to people who are potentially outside of their own network (Leek et al., 2017). In the scope of this research, the definition of engagement in two-way communication is defined by *retweets*, and the number of *retweets* is its measurement.

3.1.3 Revised Model

To reiterate this research concept is inspired by Leek's article but to be clear it has been modified to fit our research goals. Figure 1 shows the revised model with four primary stages:
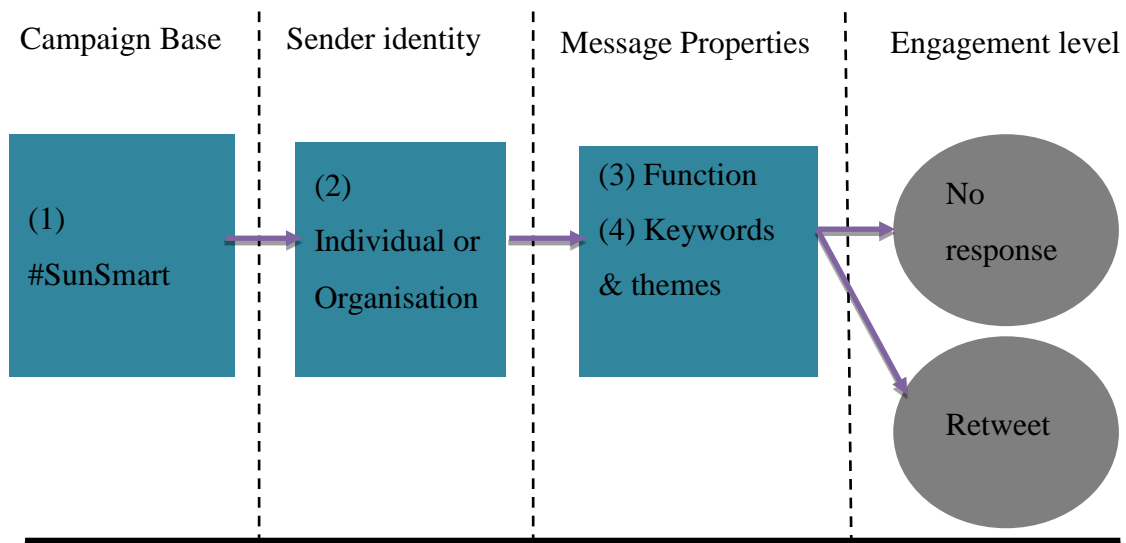
Stage 1: Data collection. Unlike Leek's study focusing on four companies, this research goes beyond the use of only one company's database. To reach out to broader audiences, regular users who may have ever discussed SunSmart but may not follow SunSmart's Twitter official account play a critical role. Therefore, this study targets all users who have ever used the hashtag "#Sunsmart." Using the hashtag #Sunsmart, a complete data set of tweets concerning the SunSmart campaign have been collected. The data is from Twitter datagrant project approved by the Ethics Committee and our data access has been authorised by the Ethics Committee in April 2018.

Stage 2: Identification of user identities & comparison levels of engagement per user types. To address the first research question (RQ1) whether the levels of engagement differ by individuals and organisations; this study further identifies the population as organisations or individuals since they are distinctive communicators and usually have different purposes for their communications and preference. With a focus on retweets, this study measures the levels of engagement by the number of retweets in descriptive statistics.

Stage 3: Explore the levels of engagement per message functions and types of users. In reply to the second research question (RQ2), understanding the primary functions that tweets serve with the public is the starting point (Lovejoy & Saxton, 2012). By using Natural Language Processing to scrutinise data, we identified four message function categories from data and developed a "codebook" for data categorisation. The codebook development was based on *all tweets* in order to look at all communication. The coding rules and definitions of categories were written in the codebook. Next, by content analysis, each tweet was manually categorised into different categories according to the criteria found in our codebook. For ensuring the coding achieved a reliable level, another coder used a sample of data to categorise per the codebook. After that, the usage of message functions by individuals and organisations were compared in descriptive statistics. Here is focused on the *retweets.*

Stage 4: Keywords and thematic topics analyses. In the final phase, TF-IDF, one of the techniques in Natural Language Processing for the keywords (RQ3a) and an hashtag visualisation thematic analysis for identifying themes (RQ3b) have been adopted to conduct an analysis among message functions to explore what makes for a *retweets.*

Figure 1. Research scope and focus of this study

Campaign Base	Sender identity	Message Properties	Engagement level

(1)
#SunSmart

(2)
Individual or
Organisation

(3) Function
(4) Keywords
& themes

No
response

Retweet

Methodology includes "Original tweets & Retweets"; Result analysis will focus on "Retweet"

In this study, the methodology serves a broader scope than our result analysis. To have an overall insight into the communication produced during the campaign, we use all data including original tweets and retweets to identify message functions, develop the codebook.

However, since the focus of this research is the level of engagement that we measured by the number of retweets. Among all these tweets, we looked which functions are the ones most retweeted. Thus, the result analysis in the following chapter will be reported by a comparison between individuals and organisations in *retweets*.

3.2 Research Scope & Data Collection (Stage 1)

This paper aims to understand how public communicate on Twitter during the SunSmart campaign by identifying the message functions and to what extent the message functions generate public engagement in the campaign. Furthermore, exploring the content of message functions and the use of language. SunSmart, the representational and worldwide skin health Population-wide health campaigns have been implemented in Australia for the past 35 years (Shih, Carter, Heward, & Sinclair, 2017), is suitable as a research target. This study is conducted from a campaign basis and collected data based on hashtag. The data is from the Twitter datagrant project, for which approval from the Ethics Committee was received. Our data access has been authorised by the Ethics Committee in April 2018. With different hashtags

associated with the SunSmart campaign, a complete dataset of tweets comprising 11,687 tweets that have been posted on Twitter in the English language, as found in the Twitter datagrant project, were collected. Since this research has a focus on the hashtag #SunSmart, the data was then filtered by hashtag #Sunsmart in the "text" column from the 11,687 tweet dataset. In sum, a total of 2,391 tweets in the English language posted by 1397 users whomsoever used the hashtag "#Sunsmart" is then the data base for this research. The data timeframe has the period from April 2014 to December 2014. For each tweet, the following metadata is included: the message text, the date and time of the message, whom the message originated from, the users' biographies, the number of retweets it received and the location from which it was posted.

SunSmart is selected as our research subject

In order to understand how to effectively exploit social media communication to enhance the public's engagement in the subject of skin health, SunSmart, a worldwide leading campaign in skin cancer prevention and cancer detection since 1988, which is funded by the Cancer Council Victoria and the Victorian Health Promotion Foundation of Australia, has been selected as the subject of this research study (SunSmart, 2018).

This campaign has been chosen for two main reasons. First of all, its leading position in the skin health industry in Australia makes SunSmart a representative candidate. SunSmart aims to improve skin cancer prevention awareness, knowledge, attitudes, and behaviour in priority populations as well as support target populations to detect skin cancers in the early stage. With the mission of reducing skin cancer incidence, morbidity, and mortality, this campaign has turned out to be a multi-faceted campaign by providing leadership and innovation in ultraviolet radiation (UV) protection and is operated in several territories of Australia by respective Councils. Second, SunSmart's campaign is relatively active on social media platforms. With the help of social media, it has reached countless people in many different countries (SunSmart, 2018). Therefore, SunSmart's campaign is an ideal research subject for conducting a study on the basis of social media platform information.

3.3 Identification of the type of users & Comparison of their levels of engagement (Stage 2)

This research is comprised of two coding parts: (a) marked message of senders' identity (individual or organisation for a tweet to investigate the degree of engagement by different types of users), and (b) interpreted message functions of a tweet.

For addressing research question one (RQ1), each tweet was coded with one of the user types: 'organisation' or 'individual' by looking at the user's profile description. Organisation' means a Twitter user is a group, which includes different types of groups, such as community (e.g., Liverpool Health Community), a foundation, or a business unit (e.g., an online shop). All of these were labelled as an organisation as defined in Appendix B. The identities of each user can very often be distinguished from their profiles, where it is usually revealing information about their activities and occupations. For example, a user's bio profile "*The Association of Chartered Physiotherapists in Oncology and Palliative Care*" was interpreted as an organisation.

In some cases, users' biographies offer only vague information such that it was unable to be identified. This might be because the user changed their content or did not provide any information. In that case, we used the user's Twitter ID number to trace their account and further identify them from their photos or published messages. If a sender's identity was still unable to be distinguished in the end, it was excluded from our analysis. After identifying user types, 5 user accounts without bio information could not be identified by other information and were excluded from the data base, this reduce my analysis database from 2391 to 2381.

Lastly, a descriptive statistics was used to compare the levels of engagement among individuals and organisations. In addition, a Chi-square test which was used to determine if there are differences in the distribution of results for each group was adopted to examine if the differences are statistically significant.

3.4 Examination of the levels of engagement per message functions (Stage 3)

3.4.1 Defining the message functions categories to develop a codebook

Before classifying tweets into different types, a codebook needed to be developed that can be seen as a guideline to categorise tweets. A codebook usually includes the rules of coding and the definitions of different categories. The first step to develop a codebook was to define the coding categories and their dimensions. In this study, the codebook development uses a more inductive method by the content analysis through data. As shown in Table 1, the four primary categories and sub-categories were identified; the detailed procedure is illustrated as follows (See sub-sections A through C):

Table 1. Primary and sub-categories of message functions

Message Function		Dimension (sub-category)
1	Functional	1.1 General problem-solving
		1.2 Suggestion/ Advice
		1.3 Disease-associated
2	Informational	2.1 Event
		2.2 Personal experience
		2.3 Opinion/ Preference
3	Interactive	3.1 Mention ("@")
		3.2 Open-questions
		3.3 Thanking
4	Promotional	4.1 Participation
		4.2 Incentive

A. Development of general categories for the codebook from literature

As seen on Table 1, the initial three primary coding categories are *Informational, Interactive* and *Functional*:

Informational and *Interactive* have been developed based on ideas from previous health communication studies. From the literature, a range of studies has shown that there are two primary functions for media usage: *information-sharing* and *dialogue* (Leek et al., 2016; Lovejoy & Saxton, 2012; Lovejoy et al., 2012; Neiger et al., 2013a; Park et al., 2013). Although the exact definition in different papers are somehow different, the basic concept is similar, therefore firstly we defines the *Informational* and *Interactive* as the two primary categories to address the purposes of messages: information-sharing and interactive.

In addition to these two, Lovejoy and Saxton (2012) state that useful information can serve the purposes of spreading information and creating dialogue (Lovejoy & Saxton, 2012). Public engagement in a health campaign may be increased when people acknowledge the value generated from the health information. Hence, the *Functional* category that is expected to emerge from health communication messages has also been chosen for this study's codebook addressing the purposes of messages: offering a functional message as a category likely solves this problem.

B. Definition of the sub-categories of codebook from the content of tweets

Computational text analysis was employed: Natural Language Processing using Python

A further scrutinising of data via computational text analysis was conducted thereafter for several reasons as is explained hereafter. First, messages passed on by data are critical. It can be imagined that prior categories from research may not be able to be completely suitable for different situations and sub-categories can give us what categories fit better in this study. Second, determining categories based merely on the prior studies might lead to some overlap or overlooked items i.e., codes included in one category also seem to be a fit for another category; Sub-categories are advantageous to enable us to know whether there is a need to reconsider the primary categories (Erlingsson, & Brysiewicz, 2017), and data scrutiny can offer an overall insight to decide the sub-categories. Third, as mentioned in 3.1.1, manual quantitative content analysis is increasing problematic when attempting to sort out the unstructured data (Waldherr, et al., 2016); a computational text analysis can enhance the efficiency on defining categories.

Thereby, after preliminarily defining the three primary categories, this researcher then scrutinized data thoroughly via the data mining software, Python, to find out the *most common words* and analyse those words' function. This enables an observation to be made regarding how many essential message functions are involved and provides us with a hint about how to determine the final primary categories and their dimensions.

A computational text analysis approach, *Natural Language Processing* (NLP) has been discussed for analysing unstructured data (Feldman & Sanger, 2006; Heyer, Quasthoff, & Wittig, 2006; Waldherr et al, 2016). NLP text mining can automatically capture semantics of texts in unstructured corpora. (Waldherr, et al., 2016) and it can be conducted via toolkit in Python. Python is widely used for data mining; it is used as a tool to conduct text analysis and extensively support diverse tasks of linguistic processing.

The Natural Language Toolkit (NLTK), which is a library suite for *natural language processing* written in the Python programming language, has been adopted. A wide range of tasks for data mining can be performed via the NLTK. In traditional natural language processing, the procedure begins with *noisy cleaning, tokenize, stemming, tagging* and the *removal of stop words* and *punctuation* for data preparation. Then, the most common words can be extracted. In this study, since grammatical rules were not in line with the research direction, *tagging* was not included in the process. Thus:

- *Eliminating noise*: Some Unicode characters (e.g., b', \x85, \x92, \x93) included in the downloaded data may lead to confusing results. Thus, removing the noise is the starting point to avoid Unicode characters expelling other meaningful words.

- *Tokenizing* refers to a process of splitting each element of text into a word. After cleaning the data noise, tokenization plays an important role to turn a string of text into each element, and this is a preparation for extracting the unigram.

- *Stemming* refers to the process of suffix removal from derived words to identify the word stem of related words.

- *Removal of stop words and punctuation*: Some stop words (e.g., the, is, was) and punctuation with high frequency are not meaningful. To find out the most powerful words, Python NLTK library embedded function can detect punctuation and stop words and remove them.

Through the analysis, the content of each tweet can be recognised as a list of words to count. For instance, the original text of a tweet: *"b' Speed up your morning beauty routine with new Cancer Council cosmetics! #SunSmart + clear skin http://t.co/XWwzmxrgDk http://t.co/NQjTO50Dfi'"* is processed as: **"'speed', 'morning', 'beauty', 'routine', 'new', 'cancer', 'council', 'cosmetics', 'clear', 'skin'"**. In this way, the words of all tweets can be counted to find out what are the most common words.

Figure 2. Text pre-processing procedure for content analysis used for this study
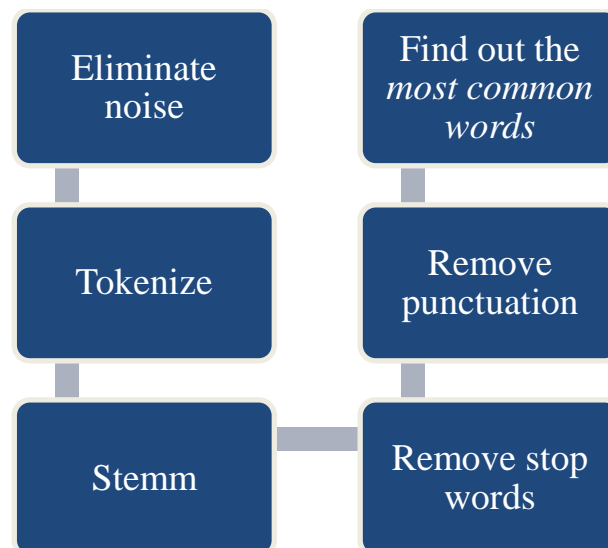
Therefore, the results in Table 2 show the most common words that were extracted by the analysis procedure as stated prior.

Table 2. Most common words and hashtags from the SunSmart dataset

Most common 20 words	Most common 20 hashtags
('sun', 400),	('#skincancer', 183),
('skin', 233),	('#summer', 146),
('sunscreen', 207),	('#sunscreen', 141),
('today', 179),	('#melanoma', 128),
('cancer', 133),	('#win', 116),
('stay', 129),	('#skincare', 90),
('enter', 123),	('#sunsmarts', 80),
('forget', 120),	('#dailyshade', 46),
('get', 113),	('#sunprotection', 41),
('day', 110),	('#health', 39),
('spf', 110),	('#sun', 38),
('tips', 102),	('#cancer', 38),
('know', 100),	('#sunsafety', 37),
('summer', 97),	('#nofilter', 35),
('see', 94),	('#maldives', 34),
('wear', 94),	('#sunscreenbands', 33),
('hat', 94),	('#giveaway', 32),
('check', 91),	('#auspol', 32),
('easy', 91),	('#sunsafe', 31)
('made', 91)	('#spf', 31)

C. Determination of the final categories & dimensions to develop a codebook as a guideline for tweets categorisation

While three primary categories were pre-defined based on categories used in prior research as described prior in 3.4.1, Table 2 results gave us a clear direction as to decide the final primary categories and distinctive sub-categories and these details are explained as follows:

(1) Determination of the sub-categories "1.1 General problem-solving", "1.2 Suggestion/ Advice", and "1.3 Disease-associated" of the *Functional* type: Table 2 shows a set of wording related to the functional type yet somehow different, this allows us to decide the sub-categories of *Functional* category. For instance, the "stay" "forget" with strong tone tended to be suggestions; "cancer"; "#melanoma" were more concerned with disease; "skin" "spf" "#sunprotection" seemed to be general solutions for sunblock. As a result, three dimensions: 1.1 problem-solving, 1.2 suggestion, and 1.3 disease-associated (Table 1) were identified through the data.

(2) Determination of the sub-categories "2.1 Event" and "2.3 Opinion/ Preference" of the *Informational* type: When decided upon as a sub-category of the *Informational* type, a pervasive dimension "2.1 Event" was intuitionally selected. After that, according to the data from Table 2, "today" is the most likely word related to that Informational type. The sub-category "2.3 Opinion" was decided after looking into the tweets containing "today."

- A tweet with words "today": *"8 years ago today, my world collapsed when I was diagnosed with skin #Cancer - **today** my lovely Mum is enjoying her birthday! #sunsmart"*

(3) Determination of the sub-categories *"3.1 Mention (@)"* of the *Interactive* type: The mention "@ " can be used in any position in a sentence. It is usually used to flag users who would like to be referred, which is a distinct feature that allowing users to deliver certain messages, republish or tag topics. Therefore, it was selected as a dimension of *Interactive*.

(4) Added one primary category "4.Promotional" and sub-categories "4.1 Participation" and "4.2 Incentive" (Table 1): Table 2 results show that some of the most common words in our data such as "enter" and "#giveaway", these words reveal an entirely different purpose: "promotion." Moving forward, because "#giveaway" and "enter" are also two different types of purposes of message, "#giveaway" implies an incentive, and "enter" shows a strong intention to encourage people to participate in activities. By selecting a couple of tweets with words "#giveaway", "enter" in the dataset and reviewed the content, the distinctive functions were confirmed. Therefore, one primary category "4. Promotional" and two sub-categories "4.1 Incentive" and "4.2 Participation" were added (Table 1).

(5) The other sub-categories "2.2 Personal experience" "3.1 Open-questions" and "3.3 Thanking" have been added in the process of qualitative review: While the top common words enabled a determination of most of the sub-categories, dimensions: "2.2 Personal experience" "3.1 Open-questions" "3.3 Thanking" (Table 1) with small data were unable to be detected from the Table 2. For these two categories, they were determined when a tweet were unable to be classified into to one of the defined categories. For the complete definitions, please refer to the complete codebook in Appendix B.

3.4.2. Manual classification of tweets & Coding Reliability test to ensure classified results reliable

After determining the definitions and code rules for message functions, a completed codebook was developed. Each tweet served as a coded unit in this study and was assigned to one of the defined four primary categories. Content analysis refers to "a research method that uses a set of procedures to make valid inferences from the text" (Weber, 1990, p.9). We interpreted the primary function and labelled each tweet with a sole category. In case a tweet appeared to serve dual purposes, codes were assigned to what was considered the primary purpose of the tweet message.

In qualitative analysis, a calculation of coding reliability is required. Here, a reliability test was conducted to confirm the coding reliability. Coding reliability is calculated from a random set of tweets (n=250) which reflects over 10% of the entire dataset (n=2381). Another human coder followed the codebook to interpret each tweet and assigned tweets into associated categories for message function and types of users independently. By using the Cohen kappa test in SPSS, a test for checking whether the judgement between the coder and the author is consistent, coding reliability is established. For this study the test results reflect a high level of coding reliability with kappa value = 0.848 and 0. 819 for types of user and message function; the coefficient was statistically significant (p<0.005).

3.4.3 Comparison of the levels of engagement in terms of message functions and types of users

After identifying the message function types, a descriptive statistics analysis is used to compare the levels of engagement among message functions and types of users. In addition, a Chi-square test which was used to determine if there are differences in the distribution of results for each group was adopted to examine if the differences are statistically significant.

3.5 Keywords and Thematic topics analysis through computational text analysis tools (Stage 4)

Two computational text analyses- Natural Language Processing and hashtag thematic topic analysis were conducted to find out keywords and themes among message functions; keywords and thematic topics were chosen due to their uniqueness to explore message composition.

3.5.1 Division of spreadsheets prior to using computational text analysis tools

In order to understand whether critical words and thematic topics differ in terms of message functions, types of tweets and type of users, the dataset is divided into a number of spreadsheets per message functions with mentioned criteria. The analysis begins with a general comparison by message functions to offer an overall picture of what keywords and thematic topics constitute each category of message function. After that, a further analysis was conducted to explore significant words and themes emerged in original tweets (i.e., post) and retweets respectively. From the comparison between messages by organisations and individuals, a more in-depth insight can be gained in the final results. Results analysis focuses on *retweets* and results for original tweets are placed in Appendices C-H for reference.

3.5.2 Keywords: TF-IDF analysis (Term Frequency- Inverse Document Frequency)

Keywords analysis is conducted by TF-IDF to identify the most significant words in a text. TF-IDF is a widely used method in text mining and information retrieval, which is used to create feature vectors and identify exclusively features in a class (Nassirtoussi, Aghabozorgi, Wah, & Ngo, 2014; Ramos, 2003). As mentioned in 3.1.1, it is one of the techniques of NLP and can be conducted via software Python as well. Many studies (Groth & Muntermann, 2011; Hagenau, Liebmann & Neumann, 2013; Luo, Zeng & Duan, 2016; Niu, Zhu, Pang, & El Saddik, 2016; Peramunetilleke & Wong, 2002) adopt this method of data classification and training and marketing area (Netzer, Feldman, Goldenberg & Fresko, 2012).Therefore, keywords are selected based on their TF-IDF score can address the attributes of a specific message function corpus to be seen as keywords.

TF-IDF stands for 'term frequency-inverse document frequency,' which reflects how important a word is to a document in a set of documents (also known as corpus or collection). The TF-IDF score calculated for a term is comprised of two parts. Term Frequency (TF) refers to a count of how many times a word occurs in a given document, and Inverse Document Frequency (IDF) is a logarithm of the number of total documents divided by the number of documents containing the term. The general idea is that TF measures the importance of a term in a document and IDF with the concept of how informative a document to the corpus. (Ghosh, & Desarkar, 2018).

In short, the TF-IDF value increases with the number of times a word appears in a document, but its weighting lowers when the word is frequently used in many documents (i.e.,

corpus). That can be seen as a measure of information amount provided by a keyword in a corpus. Mathematically, the formula of term x within a document y is denoted as:

$$TF - IDF_{x,y} = tf_{x,y} \times \log\left(\frac{N}{df_x}\right)$$

$tf_{x,y}$ = frequency of x in y
df_x = number of documents including x
N = total count of documents (i.e. corpus)

This task can also be accomplished by using NLTK programming language in the Python.

3.5.3 Thematic topics: Hashtags Visualization Thematic Analysis

The Hashtag ("#") is a unique text feature that users apply to pass on critical information on subjects. Hashtags are frequently used in conversation creation and participation. An analysis of 74 million tweets indicates that used hashtags tend to be retweeted (Suh et al., 2010). While a single hashtag standing for a topic can be used as the search string to discover a relevant discussion for a particular topic, multiple hashtags placed together can form a prominent theme, which contributes a valuable insight into what themes people mention per different message functions. Thus, the concept of thematic topics is observing the relationships between co-occurring hashtags in each message function, which is called 'hashtags visualization thematic analysis.'

To conduct a hashtags visualization analysis, a textual corpora analysis tool, Cortext Manager[5], is used, and a direct measure approach was selected for not affecting the full network map layout, which calculates the raw co-occurrence times between terms (Cortext Manager Documentation, 2017). In addition, for the purpose of getting rid of insignificant edges, I filtered top edges to display the most weighted edges according to the online guidance of Cortext Manager. Co-occurrence means that two hashtags appear in the same scope. The underlying assumption of co-occurrence is that a relationship exists between terms occurring together (Rebholz-Schuhman et al., 2007). With the help of an automated tool, a dominant thematic topics' pattern is sketched and an in-depth sense of what topics were discussed and popular in different message functions can be gained.

The practical procedure starts with extracting the words after the hashtag character (#) in Cortext Manager. Next, a corpus is established in ranking principle by frequency. After that,

---

[5] Cortext Manager. https://www.cortext.net/projects/cortext-manager/

a corresponding network graph can be drawn by conducting a mapping script, and the thematic relations can be observed. More specifically, in the network mapping a hashtag is a *node*, and two hashtags appear jointly within a single tweet are linked to be an *edge.* For instance, if "#sunsmart" and "#skincancer" occur in a single tweet, a tie between these two hashtags is established. Nodes and edges connecting each other can form a network and be visualised as a graph. The network is used to understand the structure of relations between the nodes constituting a network (Hanneman & Riddle, 2005).

The visualization analysis graphs are constructed by measuring closeness in a force-directed layout, the general concept is two nodes that have stronger connections are presented in close proximity, and the weaker connections are in the far distance. The weight of nodes is calculated based on the hashtag's occurrence in the entire dataset, and the weight of edges is calculated based on the co-occurrence of paired hashtags. The more often paired hashtags occur together, the more weighted and visible the edge is, specifically speaking, the paired topics are more frequently mentioned together. The size of nodes and edges can reveal the importance of a hashtag or a relationship.

4. RESULTS

Among a total of 2,381 tweets collected for analysis, 1,390 senders (sending users) were involved, which consists of 739 individual users and 651 organisations. In general, over half of the tweets (56.07%, n=1335) come from organisations; the proportion is slightly higher than tweets from individual users (43.93%, n=1046). Among these tweets, approximately 35.32 % of the tweets appear to be *Retweet* (n=841). As explained in '3.1.3 Revised model' chapter, on the basis of comparison between individuals and organisations, following result analyses focus on the *Retweet*.

4.1 Levels of engagement differ between organisations and individuals (RQ1)

The data on Table 3 indicates that regarding the levels of engagement (i.e., retweets) by user types, it is found that the number of reposts from personal users (52.20%, n=439) is higher than the number of reposts by organisations (47.80%, n=402). These differences in number of retweets between the user groups 'individuals' and 'organisations' are statistically significant $(X^2(1)=36.092, p< 0.0.1)$.

Table 3. Distributions of Retweets－individuals and organisations (N=841)

Sender Identity	Level of engagement (*Retweets*)
Individuals	439 (52.20%)
Organisations	402 (47.80%)
Total Tweets	841

This result shows that levels of engagement differ between individuals and organisations; and individuals engaged more in the SunSmart campaign than organisations. This result may imply that individual users are more interested in sharing information regarding SunSmart campaign than organisations. This could also be the reason that *Retweet* is not the primary mechanism for organisations to engage in the SunSmart campaign.

4.2 Levels of engagement differ per different message functions and types of users (RQ2)

A further analysis sought to determine whether the levels of engagement differ per different message functions and types of users.

4.2.1 Levels of engagement differ per different message functions

Regarding the levels of engagement (i.e., retweets) among message functions, the data on Table 4 shows that the *Interactive* category comprises the most retweets (n=401), followed by *Functional* (n=305) and *Informational* (n=98), while *Promotional* category comprises a low number of retweets (n=37).

After investigation in detail regarding the levels of engagement (i.e., Retweets) among message functions, the result shows that level of engagement differ among message functions. The difference in numbers of retweets per message function are statistically significant $(X^2(3)=28.007, p< 0.0.1)$. This result might imply that different message functions lead public to the different level of engagement in the health campaign. The finding is also supported by the results of Leek's paper in which they suggest that tweet functions and the language used in the content of messages affect behavioural engagement (Leek et al., 2017).

Table 4. Distribution of Retweets by message function (n=841)

Message Function category	Level of engagement (*Retweets*)
Functional	305 (42.90%)
Informational	98 (28.99%)
Interactive	401 (33.25%)
Promotional	37 (29.37%)
Total of Tweets	841

4.2.2 Among different message functions, the levels of engagement in regard to Retweets differ between individuals and organisations

Analysing these results from the perspective of users' identity–individual vs. organisation, we can gain a further understanding of how the two groups are engaged in the SunSmart campaign per different message functions. In regard to the levels of engagement per each message function, Figure 3 illustrates that individuals engaged in **Functional** (n=159) messages slightly more than organisations (n=146). *Functional* type is defined as offering suggestions or information that might solve problems. The result show that individual users are more concerned with information which can offer a solution to a health problem than organisations. This tendency may imply that the individuals are inclined to share information as long as the usefulness of a message is clearly evident.

However, for **Interactive** type, individuals have a slightly lower level of engagement in *Interactive* messages (n=196) than organisations (n=205). In our categorised data, the

Interactive category mainly consists of the sub-category "mention (@)," this might imply that tweets including "mention (@)" could maximise the ease of initiating with organisations than individuals.

On the other hand, the higher numbers of retweets explain that individuals (n=70) are much more likely to share an ***Informational*** message than organisations (n=28), whereas personal users are seemingly not much engaged in ***Promotional*** message (n=14) than organisations (n=23). In our categorised data, the tweets of the *Informational* category mainly composed of the sub-category *"Opinions/Preference,"* this might be inferred that the individuals are more interested in sharing information about personal opinions or preference than organisations. However, the *Promotional* messages are less attracted to individuals than organisations, this could be explained that individuals may not be taking advantage of the promotional incentives or not interested in the promotional wording in comparison with organisations.

Figure 3. Distribution of *Retweets* by message function－individuals vs. organisations
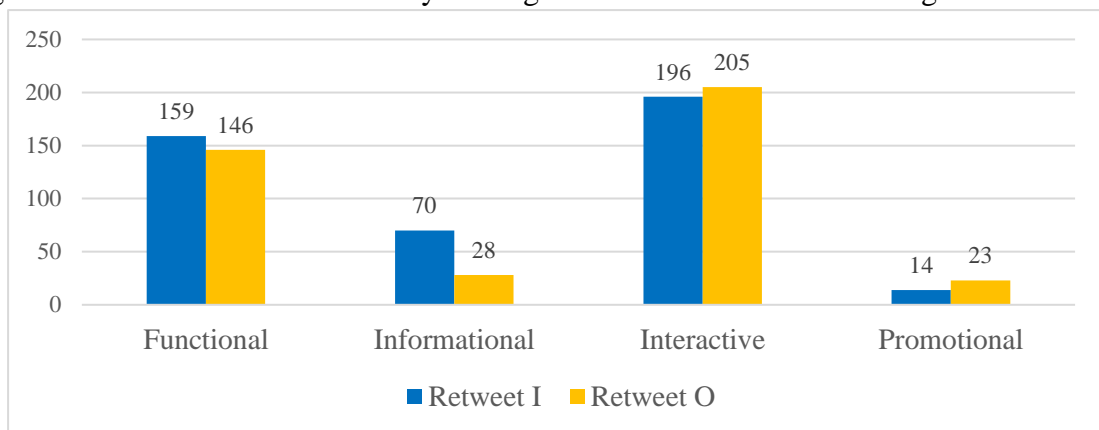
Note: Individual (I) and organisations (O)

4.3 Keywords & thematic topics differ per different message functions among individuals and organisations (RQ 3)

In order to understand whether the critical words and thematic topics differ in terms of message functions, types of users, and levels of engagement of users (i.e., retweets), the data is prepared as stated in the prior section on Methodology, 3.5.1., and using the TF-IDF method via natural language processing and hashtags thematic visualisation analysis to find out what keywords and thematic topics constitute each category of message function in *Retweet*. The results are depicted in Table 9 (Appendix C) indicating that keywords and thematic topics differ per

different message functions. The main thematic topics of the four message functions are interpreted from the hashtags visualisation analysis chart in Appendix E-H. Following analyses are further reported by a comparison between individuals and organisations in *Retweets* per message functions.

4.3.1 Keywords of tweets related to the SunSmart campaign differ in regard to message categories among individuals and organisations (RQ3a)

Result: Different keywords are used by individuals and organisations

The data on Table 5 indicates that the differences still outweigh the similarities when comparing keywords in the *Retweets* by individuals and organisations. While some mutual terms are observed in *Informational*, *Functional,* and *Interactive* messages, such as "cancer " and "protected" in the *Functional* category, the variation of keywords among message functions and types of users is still obvious. Particularly, when looking at the retweets in the *Promotional* function category; the keywords are entirely different per different user group. Individuals are more engage in the SunSmart by slogan words (e.g., slip, slap, slop) but organisations prefer to share information including encouraging words (e.g., download, join). This may imply that individuals and organisations have broadly different tastes in regard to what engages them enough to retweet a *Promotional* message.

On the other hand, the minor mutual keywords "cancer," "protected," "skincare," and "sunbeds" among different message functions may also give us a direction about how to compose a message to engage both types of users.

Table 5. Keywords of message functions in Retweets－individuals vs. organisations

Message Function	Retweets	
	Individual	Organisation
Functional	. skincare . common . protected . cancer	. dont . cancer . strong . protected
Interactive	. know . spot . look . skincare	. strong . protected . shade . skincare
Informational	. palmtrees . sunbeds . maldives . nofilter	. monday . july . allowed . sunbeds
Promotional	. doncasterisgreat . slap . slop . slip	. video . victoria . download . join

*Note: the difference between individuals and organisations are highlighted in *red*

4.3.2 The thematic topics of tweets related to the SunSmart campaign differ in regard to message categories for individuals and organisations (RQ3b)

 (1) Thematic topics in *Functional* message: Retweet

Result: While the results show that both individuals and organisations prefer "general skin health issues and skin cancer (theme 1)" most, their focus on theme 1 differs
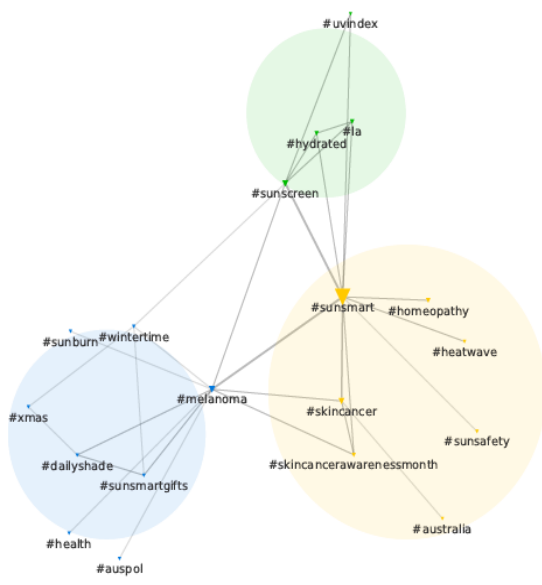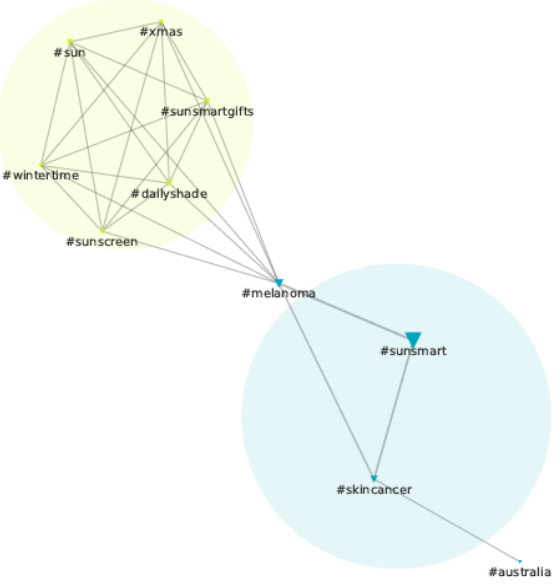
As shown by the data on Figure 4, comparing the Retweets posted by individual and organisation users, overall, both user groups have the greatest interest in forwarding *general skin health issues and skin cancer messages* (theme 1) followed by *wintertime sun-protection* (theme 3). However, *sunscreen usage t*o *avoid hydration in LA* (theme 2) is more attractive to individual users. This may impy that although both types of users care about general skin issues and wintertime protection, individual users are more interested in the sunscreen usage topics in LA area than organisations.

 Looking into the details, although both parties prefer theme 1 most, the difference are most evident in observation of the top three nodes: #sunscreen, #skincancer, and #melanoma. For individual users, the weight of these three topics are comparable. However, the pattern in

the organisation's map reveals that #melanoma and #skincancer with closer proximity are more critical than #sunscreen. From this result it could be said that individuals are more concerned about daily protection from the diseases than the diseases themselves.

Another interesting point is, #melanoma is more interconnected in scope than #skincancer or #sunscreen as seen in their respective network maps. This attribute may imply that a message comprised of multiple hashtags along with #melanoma and #sunsmart is an intriguing topic for both groups.

Figure 4. Thematic topics in the *Functional* message function of *Retweets*

Individual	**Organisation**
Theme 1:general skin health issues and skin cancer ⇨ #skincancer, #melanoma *Theme 2: sunscreen usage to avoid hydration in LA* ⇨ #sunscreen, #uvindex, hydrated, #la *Theme 3: wintertime sun protection* ⇨ #melanoma, #wintertime, #dailshade, #xmas	*Theme 1:general skin health issues and skin cancer* ⇨ #skincancer, #australia, #melanoma *Theme 3: wintertime sun protection* ⇨ #dailshade, #sunscreen, #wintertime

(2) Thematic topics in Interactive message: Retweets

Results: While both types of users prefer to share information concerning theme 1 to theme 4, the focuses and value placed upon the topics differ

The data charted on Figure 5 illustrates that both types of users generally show a broad interest in *Interactive* Retweets, but focus on different minor topics. For *activity, sun protection and skin cancer* (theme 1), individuals preferred to share information about *sports events* (i.e., #im703mandurah) in *Interactive* messages, but organisations forwarded information about *a prize-draw activity* (#win, #summer). For theme 2, individuals are interested more in topics such as #uv, #sunscreen, #melanoma but organisations preferred #health, #skinweek, #prevention, and #cure. This might be interpreted that individuals are more concerned with usage of UV protection to avoid melanoma, while the organisations' focus on diverse diseases' protection.

Another useful insight is in regard to theme 4 with hashtags #stylish and #giveaway that show a more close relationship with #sunsmart on the individual map than on organisation map, which implies that individual users have a higher interest in republishing *marketing information* (theme 4),such as distribution of gifts than organisations.

In addition, *community marketing* (theme 5) is more appealing to individuals per our results. This may also imply that community topics are more interesting to individuals than organisation.

Figure 5. Thematic topics in the *Interactive* message function of *Retweets*

Individual	Organisation
Theme 1: *activity, sun protection and skin cancer* ⇨ #sunsafety, #imwa, #im703mandurah, *Theme 2: health disease prevention* ⇨ #uv, #sunscreen, #melanoma, *Theme 3: skin cancer* ⇨ #skincancer, #skincare *Theme 4: product marketing* ⇨ #stylish, #giveaway *Theme 5: community marketing* ⇨ #giveaway, 30secondmom	*Theme 1: activity, sun protection and skin cancer* ⇨ #win, #summer, #melanoma, #dailyshade, #doncasterisgreat *Theme 2*: *health disease prevention* ⇨ #health, #skinweek, #prevention #cure *Theme 3: skin cancer* ⇨ #skincancer *Theme 4: product marketing* ⇨ #giveaway, #stylish

(3) Thematic topics in Informational message: Retweets

Results: sun-protection in vacation related themes (theme 3 & theme 2) are more intriguing to individuals, while organisations prefer to forward event-related information (theme1)

The data charted on Figure 6 shows an entirely different preference can be seen in retweets within the *Informational* message function. Individual users are engaged more by sun-protection in vacation association themes (theme 3 and theme 2), and the #nofilter and #maldives hashtags are the two main nodes composing these two themes. However, organisational users preferred to share information about *event* (theme 1) than republish the other two themes. This may imply that individuals are fond of holiday information more so than organisations.

Figure 6. Thematic topics in the *Informational* message function of *Retweets*

Individual	Origanisation
Theme 3: sun protection in vacation-chemical block. ⇨ #nofilter, #maldives; #sunscreen, #sea, #summer *Theme 2: sun protection in vacation-physical block.* ⇨ #hat, #shade	*Theme 1: event-related* ⇨ #im703mandruah, #sunbeds, #schoolsurfing titles
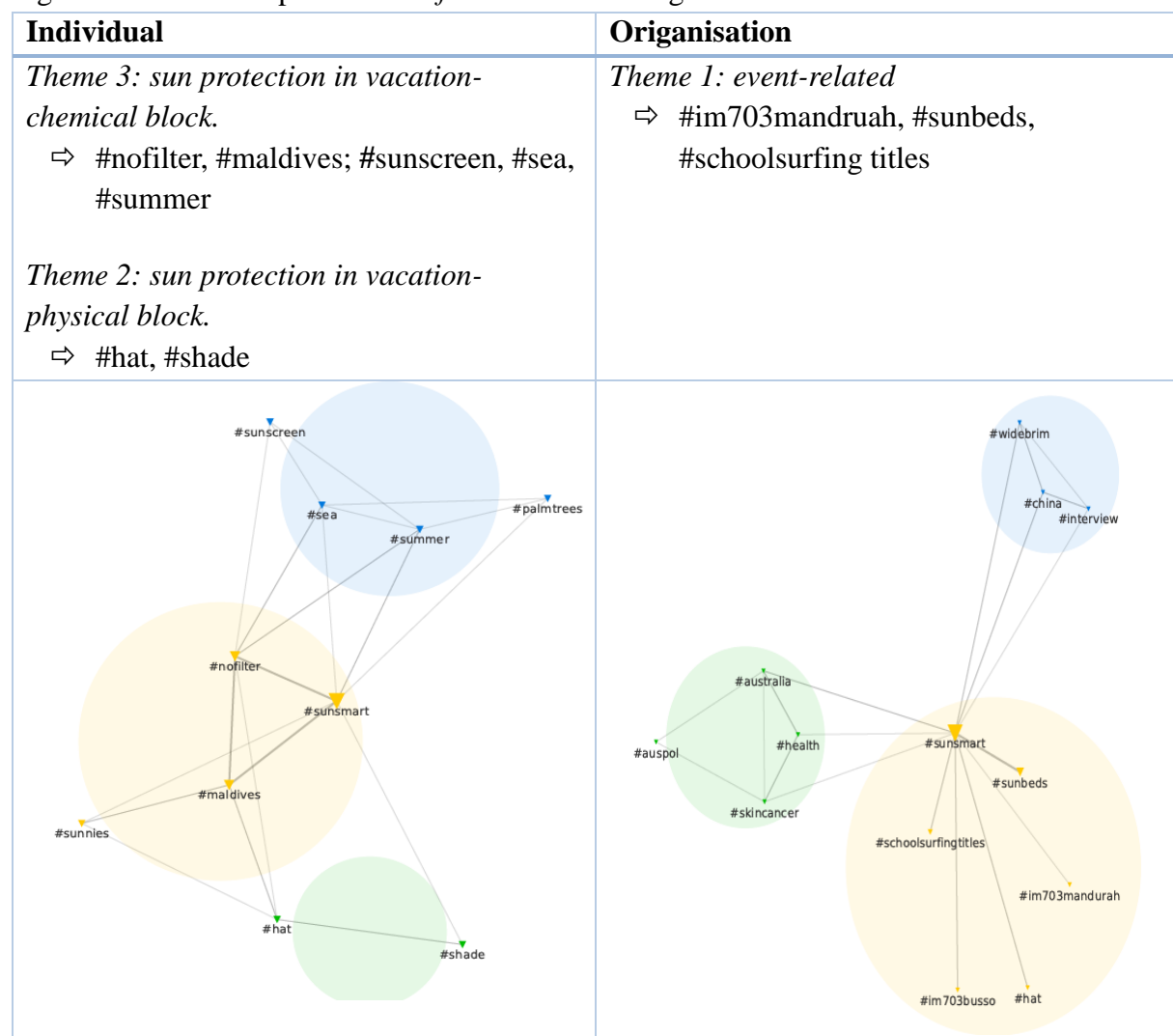
(4) Thematic topics in Promotional message: Retweets

Results: Both types of users engaged in inspiring slogan (theme 1); the difference is that individuals prefer the sunscreen and promotion (theme 3) but organisations were more engaged by the bamboo product-promotion (theme 2)

According to the Figure 7, it is found that it is found that the *inspiring slogan* (theme 1) matters. Also, we can see that both types of users flagged a location in regard to theme 1 (#doncasterisgreat, #victoria), this may imply that location is an important component for creating an intriguing *Promotional* message.

On the other hand, individual users and organisations have different preferences in regard to product promotion, the former are more engaged in the *sunscreenband promotion* (theme 3), but the latter was more interested in *bamboo product promotion* (theme 2).

Figure 7. Thematic topics in the *Promotional* message function in *Retweets*

Individual	Organisation
Theme 1: *inspiring slogan* ⇨ #doncasterisgreat *Theme 3: sunscreenband promotion* ⇨ #sunscreenband, #giveaways	*Theme 1*: *inspiring slogan* ⇨ #doncasterisgreat, #victoria, *Theme 2: bamboo product promotion* ⇨ #bamboo, #upf, #melanoma

In summary, Table 6 shows a comparison of thematic topics among different groups, individuals vs. organisations. Any obvious difference in theme preference is highlighted in *red*. A theme shown among both user groups but which apparently has a different focus on minor topics is marked with star (*). Overall, a high variety is observed.

Table 6. Themes of message function in Retweets－individuals vs. organisation

Message function	Main thematic topics	Themes in Retweets	
		Individual	Organisation
Functional	Theme 1: *general skin health issues and skin cancer* Theme 2: *sunscreen usage to avoid hydration in LA* Theme 3: *wintertime sun protection* Theme 4: *sun protection in Melbourne*	▪ Theme 1* ▪ Theme 2 ▪ Theme 3	▪ Theme 1* ▪ Theme 3
Interactive	Theme 1: *activity, sun protection* Theme 2: *health disease prevention* Theme 3: *skin cancer* Theme 4: *product marketing* Theme 5: *community marketing*	▪ Theme 1* ▪ Theme 2* ▪ Theme 3 ▪ Theme 4 ▪ Theme 5	▪ Theme 1* ▪ Theme 2* ▪ Theme 3 ▪ Theme 4
Informational	Theme 1: *event-related* Theme 2: *sun protection in vacation-physical block.* Theme 3: *sun protection in vacation-chemical block.*	▪ Theme 3 ▪ Theme 2	▪ Theme 1
Promotional	Theme 1: *inspiring slogan* Theme 2: *bamboo product promotion* Theme 3: *sunscreenband promotion* Theme 4*: a football activity in Montclair*	▪ Theme 1 ▪ Theme 3	▪ Theme 1 ▪ Theme 2

Note (1): the difference between individuals and organisations are highlighted in *red*
Note (2): * means the theme is the same, whereas the focused upon minor topics apparently differ

5. KEY FINDINGS & DISCUSSION

5.1 Summary of Key Findings

In regard to the comparative analysis from the perspective of the types of users, individuals vs. organisations, some key findings are found which are highlighted in the following.

Finding 1: Individual users are more engaged in the SunSmart campaign than organisations (RQ1)

Regarding the RQ1 whether levels of engagement differ by organisations and individuals, the results show that individuals have a higher level of engagement than organisations.

Finding 2: Health communications of the SunSmart campaign can be categorised into messages with four main purposes and the levels of engagement per message functions and types of users differ (RQ2)

Regarding the question to what extent do message functions observed from the manifest content of tweets differ and do the levels of engagement differ per the use of the message functions and types of users, the answer is positive.

Results show that tweets sent during the SunSmart campaign can be categorised into four types of messages based on a specific function (i.e., Functional, Informational, Interactive, and Promotional). Among different message functions, the data charted on Table 4 (Chapter 4) shows a difference in engagement per different message functions. The *Interactive* category has the highest level of engagement, followed by *Functional*, and *Informational*, while *Promotional* category has a considerably low level of engagement.

Comparing the levels of engagement by individuals and organisations via *Retweet*, the data charted on Figure 3 (Chapter 4) shows a difference in engagement per different message functions by types of users. For the *Functional* messages concerning the SunSmart campaign, individuals have a slightly higher level of engagement than organisation. Conversely, for *Interactive* messages, organisations have a slightly higher engagement than individuals instead. On the other hand, individual users prefer to share *Information* message than organisation, whereas they are not fond of sharing *Promotional* types of message than organisations.

Finding 3: Individuals and organisations use different keywords and thematic topics according to the message function (RQ3a & RQ3b)

Regarding the RQ3 whether keywords and thematic topics differ in different message functions and types of users, from the results of the keywords analysis, we found that except for "cancer ", "protected", "skincare", and "sunbeds" are the four mutual keywords used by both individuals and organisations, a high variation of keywords among the four categories is observed. The most notable difference show in *Promotional* messages in which individuals and organisations use entirely different keywords. Similarly, for thematic topics, even though the individuals and organisations use a few same themes in message functions, the focus (i.e., minor topics) are different. Overall, as summarised in Chapter 4 (Table 5 & Table 6), the differences still outweigh the similarities.

To sum up, from the findings of this study, a variation in the content of tweets is found in terms of keywords and thematic topics per message function leading to differing levels of engagement for individuals and organisations.

5.2 Discussion

Despite the fact that Twitter is a widely used social media platform for health campaign information dissemination, how to strategically use Twitter to communicate with users remains unclear (Lovejoy et al., 2012). In this study, regarding how are different types of tweets are used for communication during the SunSmart campaign and resulting in public engagement on Twitter were answered. By exploring what message functions can result in higher engagement and understanding how message content play out among the functions, the results of this study offer insights how to strategically use tweets for communication and show a more explicit direction to construct effective messages.

Discussion 1

This research started by looking at how the different types of users engage in the SunSmart health campaign on Twitter (RQ1). By analysing users' profiles to identify user types and a statistical analysis to measure the public engagement, we found that individuals are more engaged in the campaign than originations, this may imply that individuals are more interested in the SunSmart campaign, and more fond of sharing information regarding the campaign than organisations. Another possible factor could be that organisations prefer to use other

mechanism (e.g., Like, Comment) instead of *Retweet* to engage in the SunSmart campaign on Twitter.

Discussion 2

Furthermore, to identify what the function of tweets (i.e., purpose of message) that individuals and organisations use to communicate and investigate how they generate engagement (RQ2), a computational text analysis was conducted utilising Natural Language Processing to develop a codebook and demonstrate that four main message functions are prominent in the SunSmart campaign messaging data. Thereafter, a qualitative content analysis was conducted to manually categorise tweets into the four message functions with a high, inter-coder reliability back up the finding. From these four message functions: *Functional, Interactive, Informational,* and *Promotional,* it is found that there are indeed different levels of engagement (i.e., retweet). By using the statistical analysis, a more detailed comparison per user type via retweets, we found that a difference in engagement among different message functions. The finding is supported by Leek's paper in which showing that the message function and linguistic style used in the content of messages can affect behavioural engagement in terms of *Likes, Retweets,* and *Comments.* (Leek et al., 2017).

Analysing these results from the perspective of users' identity–individual vs. organisation, it is found that individuals have a slightly higher level of engagement than organisations in **Functional** messages. Because the *Functional* type is defined as offering suggestions or information that might solve problems. This tendency may imply that the individuals are inclined to share information as long as the usefulness of a message is clearly evident. Conversely, for **Interactive** messages, organisations have a slightly higher engagement than individuals instead. In our categorised data, the tweets of the *Interactive* category mainly consist of the sub-category "mention (@)", this might imply that using *Interactive* tweets including "mention function (@)" could maximise the ease of initiating interaction with organisations, and increase the likelihood of organisations' engagement in the SunSmart campaign.

On the other hand, of the tweets analysed, much fewer tweets were categorised as *Informational* and *Promotional*, in which a difference also exists between these two types of users. It is noticed that individuals have a higher level of engagement in *Informational* message than organisations, but they have fewer interests in *Promotional* message types. In our classified data, **Informational** category includes more tweets regarding personal opinions or

preferences. The results may explain that organisations are less interested in retweeting such messages than individuals for the information-sharing purpose. Besides, ***Promotional*** message type is less attracted to individuals, indicating that individuals may not be taking advantage of the promotional incentives or not considering the messages as high value to share.

Discussion 3

Last, we look at the content and use of language among different message functions (RQ3a & RQ3b). By adopting the computational text analyses, the use of Natural Language Processing to reveal the keywords and the use of network visualisations to identify the primary thematic topics among message functions, it was found that in general keywords and thematic topic per different message function varied. The differences may imply that generally individuals and organisations have quite different tastes in regard to how they are engaged in the SunSmart campaign. On the other hand, the minor similarities also favour us insights into composing a tweet per message functions to engage both individuals and organisations.

Differences per message functions among individuals and organisations

Regarding the *keywords analysis*, overall, the variation of keywords among message functions and types of users is obvious. The difference is particularly notable in the *Promotional* function category in which individuals and organisations use entirely different keywords. Individuals are more engage in the SunSmart by slogan words (e.g., slip, slap, slop) but organisations prefer to share information including encouraging words (e.g., download, join). This could be explained that individuals and organizations have broadly different tastes in regard to what engages them enough to retweet a *Promotional* message.

For *thematic topics analysis,* the results in ***Funcitonal*** messages could be said that individuals are more concerned about daily protection (i.e., #sunscreen) from the diseases than the diseases themselves (i.e., #melanoma, #skincancer) and this is opposite to organisations. In addition, it could also be explained that individual users are more interested in the sunscreen usage topics in LA area than organisations as *sunscreen usage* t*o avoid hydration in LA* (theme 2) is more attractive to individual users.

Besides, the results in ***Interactive*** messages may be interpreted that individuals preferred to share information about *sports events* (i.e., #im703mandurah) for interaction, but organisations forwarded information about *a prize-draw activity* (i.e., #win, #summer). Also, we found that individuals might be more concerned with usage of UV protection to avoid

melanoma (i.e., #uv, #sunscreen, #melanoma), while the organisations' focus on diverse diseases' protection (i.e., #health, #skinweek, #prevention, and #cure). Moreover, individual users have a higher interest in republishing *community marketing* (i.e., #30secondmom) than organisations per our results. This may also imply that community activities are more interesting to individuals than organisation.

For the **Informational** messages, it is found that individual users are more engaged by *sun-protection in vacation* association themes, especially in the area "#maldives"; However, organisational users preferred to share information about *event* (i.e., im703mandruah). These results may imply that individuals are fond of holiday information more so than organisations.

Regarding **Promotional** messages, individual users are more engaged in the *sunscreenband promotion* (theme 3), but organisations are more interested in *bamboo product promotion* (theme 2). This could be explained that individual users have a higher interests in the sunscreenband product than organisations.

Minor similarities per message functions among individuals and organisations

On the other hand, some minor similarities may also favour us a direction for composing a tweet with a particular message function which can engage both individuals and organisations in the SunSmart campaign.

From the results of *keywords* analysis, the few mutual keywords: "cancer ", "protected", "skincare", and "sunbeds" are used in the *Functional, Interactive*, and *Informational* messages, which might imply that to some extent that these words of message functions can be utilised to engage both individuals and organisations.

Likewise, the results of *thematic topics analysis* may imply that by both types of users value *wintertime sun-protection* theme in terms of the **Functional** function. Also, we found about #melanoma is more interconnected in scope in both types of users' network maps, this attribute might indicate that a message comprised of multiple hashtags along with #melanoma and #sunsmart is an intriguing topic for both groups. For **Interactive** messages, the results may imply that using '@' mention funciton altogether with the *skin care* and *product marketing* themes are effective to create engagement for both types of users. For composing a **Promotional** message, it is found that the *inspiring slogan* matters. Besides, we can see that both types of users flagged a location (#doncasterisgreat, #victoria) in this type of message, this might be explained as that people are more appealing to *Promotional* messages for a particular area.

Discussion 4

Based on the research of the SunSmart campaign, this study has identified that keywords, thematic topics differ among message functions, types of users, and levels of engagement. These results may imply that tailoring compelling messages to enhance the public engagement in the SunSmart campaign is possible. For individuals or organisations who are dedicated to promoting the SunSmart campaigns (e.g., volunteers, skin cancer association), when they compose tweets to communicate with their followers who potentially have the same interest in health campaigns, following suggestions might be helpful.

On the one hand, if they are unable to identify their audiences as individuals or organisations, using *Functional* or *Interactive* messages may enhance audiences' engagement in the SunSmart campaign and reach out more potential audiences through retweeting. More specifically, they can use the mutual keywords (i.e., "protect" "cancer" "skincancer") and thematic topics (e.g., wintertime sun-protection) among these two message functions to construct the two types of messages.

On the other hand, if they can identify their audience as individuals or organisations, tailoring of a message for target audiences can refer to the keywords and thematic topics in *Retweets* for different message functions as well, which could allow the SunSmart campaign to become more visible through retweeting. For instance, for the purpose of offering general solutions (i.e., ***Functional***), it is suggested to compose messages about daily protection (i.e., #sunscreen) from diseases for individuals and construct messages focusing more on diseases themselves (i.e., #melanoma, #skincancer) for organisations. These examples may be useful to develop a marketing strategy in a specific manner.

As a result, a systematic path is humbly proposed for constructing an effective message. Following these steps may not only benefits health communication but also contributes to a health campaign promotion. When organisations organise a campaign in social media or individuals who aim to promote a health campaign, they need to consider:

- Step 1: Who the target audience is.
- Step 2: What the purpose of the message function is.
- Step 3: Adding corresponding keywords to the message may amplify a particular message function of a tweet.
- Step 4: Using the significant thematic topics as emerged per this study's results in regard to *Retweets* per message function to compose tweets.

5.3 Limitations & Future Work

These results should be interpreted while keeping in mind some limitations. First, this research was conducted using selected data within the study's timeframe, and it may possibly be not substantial enough to represent general trends or preferences of an entire population. For future research, analysing a dataset with a longer timeframe will help to confirm the findings of this study. Additionally, conducting a longitudinal analysis will help to observe the variation of content and engagement over time.

Second, this research is focused on data filtering with hashtag #sunsmart. However, among the SunSmart campaign datagrant project, there is also included several other important hashtags representing the SunSmart campaign, such as #melanoma and #skincancer. Future work can expand the research scope to the whole dataset of 11,687 tweets of the Twitter datagrant project.

Third, the depth of visualization analysis is limited. Although the visualization analysis tool is helpful to highlight crucial themes and deal with a large scale of data; however, our findings mainly focus on the prominent keywords and themes, and some information is ignored. Even though such loss of minor data might not influence our results, it cannot be denied that pieces of information may also favour some different insights and conclusions. This weak point also means that visualization analysis cannot completely replace qualitative analysis.

Apart from these concerns, this research focuses on the number of retweets, future work can enlarge the scope by including the "Like" and "Comment" categories of tweets as indicators of the level of engagement.

Last, this study analysis focuses on the population divided into only two general group individual and organisation, without distinction of their location. To have further insight into how different populations use Twitter as a health communication tool and what keywords and thematic topics are used to frame their tweets, narrowing down the population may yield more benefits to the health campaign. For instance, the usage of Twitter by organisations may differ in consideration of different strategic aims and variations might be due to their types (e.g., service or product companies) (Burton, & Soboleva, 2011), a comparison between the sectors of organisations can favour an in-depth sense to develop an effective communication strategy. Also, a geographical analysis would be interesting. By comparing countries or cities, we can have a better understanding of what kind of themes and message functions are more attractive to their populations in those particular areas.

5.4 Academic & Practical Contributions

Despite the limitations, this study contributes to research in three aspects.

First, this study contributes to ***social media*** and ***communication*** research by showing the potential of social media as platforms for communication processes (Hambrick et al., 2010; Park et al, 2016; Leek et al., 2017; Priante et al, 2018a) to discover Twitter users' preferences. The results of this research show that social media, specifically, Twitter, provides a platform to research how to effectively use social media.

Second, this study contributes to the area of ***health communication***. The effective health communication may influence people's behaviour, and effective communication can be achieved by enhancing the value of social media and understanding the two-way communication (Adewuyi & Adefemi, 2016). The results of this study show four dominant message functions and varied health topics that the public engaged in the case of the SunSmart campaign, which contributes to developing effective health communication strategies.

Third, this study contributes to ***marketing*** research by assisting educators to understand the use of social media in the marketing arena and use its insights to achieve or measure engagement with target audiences. This paper contributes to call for more researches that dedicate to marketing area for health campaigns by investigating how twitter message content varies per message functions can be used to drive publics' engagement in a health campaign.

Last, this study offers a novel methodology that combines the use of computational text analysis tools (i.e., natural processing of language and visualisation network analysis) with the manual content analysis, and a descriptive statistical analysis. This study demonstrates the feasibility of using a mixed method via an automated processing tool to deal with the data from social media by providing a multifaceted-approach to discover Twitter users' preferences (Bail et al., 2017; Priante et al., 2018b).

In practice, this research contributes valuable insights into the opportunities and limitations of using social media in improving public online engagement for health campaigns. First, this study shows how to strategically use social media for different communication purposes, e.g., interactively target audiences and promote campaigns. In addition, this study takes into account engagement of target audiences thus the results offer a perspective on the publics' preferences, which benefit health organisations or individuals (i.e., volunteers) who are dedicated to promoting health campaigns.

Second, exploring the semantic content of the tweets and by their use in discourse, namely by keywords and thematic topics among different message functions with differing levels of engagement thus offering a more profound insight into how to construct a valid, effective message. The SunSmart association can use keywords and thematic topics to create conversations and connect with other potential audiences who are interested in similar health concerns or interests. Effectively using social media also reduces a non-profit organisations' financial resources allocated to marketing activities (Park et al., 2013). This study offers a strategic health communication mechanism via Twitter, which may help individuals or health organisations to effectively use Twitter for health communication and tailoring of a marketing plan, and further to make the SunSmart campaign more visible through retweets.

Last but not least, these results may help the SunSmart campaign to engage a broader audience base. This study favours a perspective that is seldom offered in prior works since most research papers focus on active users (i.e., followers) as research base. Prior research results may not serve as a guide to identify regular users' preferences in regard to preferred message types and topics.

6. CONCLUSIONS

Twitter offers remarkable opportunities for social enterprises such as the SunSmart campaign to engage audiences in conversations (Lovejoy & Saxon, 2012). A set of communication functions on Twitter serves as an interactive application that can aid in users' engagement, (Lovejoy, et al., 2012) while how to use social media to create public's online engagement is still puzzled. Additionally, although social media platforms are less costly than traditional marketing channels, non-profit organisations or individuals may still lack resources (i.e., human capital, marketing budgets) to maintain and carry out marketing activities. Therefore, how to best use Twitter for health communication and promotion is essential and critical insights can be gained by exploring the content with which the public engages. This exploratory study provides insights in regard to how to strategically use tweets for communication, composing effective messages posted via Twitter and identifies further opportunities for using social media for health communication.

Managerially, this study's findings provide specific suggestions in regard to frame a Twitter message per message functions to enhance public engagement. An appropriately social media deployment enables to construct a message to fit potential customer' needs and observe feedback from audiences. The keywords and thematic topics analyses offer comprehensive and comparative insights by understanding the structures and patterns of hashtags and the most relevant words in each message function category, which provide useful suggestions for SunSmart, and other organisations or individuals, to construct and refine their communication content and deliver compelling messages rather than just disseminating information indifferently or randomly.

In conclusion, while framing tweets for health campaign, three points should be taken into account: first, the target audiences (i.e., individuals or organisations); second, the purpose of posting the tweet; and finally, choosing intriguing thematic topics for target audiences and combining keywords in line with the corresponding message functions. This approach may very likely help the SunSmart campaign, other health organisations or individuals who are dedicated to promoting the SunSmart campaign, or campaigns like it, to raising public engagement and facilitating faster sharing of information.

REFERENCE

Adewuyi, E. O., & Adefemi, K. (2016). Behavior change communication using social media: a review. *International Journal of communication and health*, *9*, 109-16.

Amina Jama Mahmud, Olander1, E., Eriksén, S. and Haglund, B.J.A. (2013) Health communication in primary health care - A case study of ICT development for health promotion. BMC Medical Informatics and Decision Making 2013, 13:17.

Ashley, C., & Tuten, T. (2015). Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement. *Psychology & Marketing*, *32*(1), 15-27.

Bail, C. A. (2016). Cultural carrying capacity: Organ donation advocacy, discursive framing, and social media engagement. *Social Science & Medicine*, *165*, 280-288.

Bail, Christopher A., Taylor W. Brown, and Marcus Mann. (2017). "Channeling Hearts and Minds: Advocacy Organizations, Cognitive-Emotional Currents, and Public Conversation." *American Sociological Review* 82(6):1188–1213.

Bortree, D. S., & Seltzer, T. (2009). Dialogic strategies and outcomes: An analysis of environmental advocacy groups' Facebook profiles. *Public relations review*, *35*(3), 317-319.

Brodie, R. J., Hollebeek, L. D., Jurić, B., & Ilić, A. (2011). Customer engagement: Conceptual domain, fundamental propositions, and implications for research. *Journal of service research*, *14*(3), 252-271.

Boyd, D., Golder, S., & Lotan, G. (2010, January). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System sciences (hicss), 2010 43rd hawaii international conference on* (pp. 1-10). IEEE.

Burton, S., & Soboleva, A. (2011). Interactive or reactive? Marketing with Twitter. *Journal of Consumer Marketing*, *28*(7), 491-499.

Bravo, C. A., & Hoffman-Goetz, L. (2017). Social media and men's health: a content analysis of Twitter conversations during the 2013 Movember campaigns in the United States, Canada, and the United Kingdom. *American journal of men's health*, *11*(6), 1627-1641.

Bruhn, M., Schoenmueller, V., & Schäfer, D. B. (2012). Are social media replacing traditional media in terms of brand equity creation?. *Management Research Review*, *35*(9), 770-790.

Campbell, B. C., & Craig, C. M. (2015). Health professions students academic and personal motivations for using social media. *International Journal of communication and health*.

Chang, Y. T., Yu, H., & Lu, H. P. (2015). Persuasive messages, popularity cohesion, and message diffusion in social media marketing. *Journal of Business Research*, *68*(4), 777-782.

Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, *5*(11), e14118.

Cho, M., Schweickart, T., & Haase, A. (2014). Public engagement with nonprofit organizations on Facebook. *Public Relations Review*, *40*(3), 565-567.

Cortext Manager Documentation (2017). Retrieved from https://docs.cortext.net/analysis-mapping-heterogeneous-networks/mapping-edges-definition/

Dolan, R., Conduit, J., Fahy, J., & Goodman, S. (2016). Social media engagement behaviour: a uses and gratifications perspective. *Journal of Strategic Marketing*, *24*(3-4), 261-277.

Donovan, R. J. (2011). The role for marketing in public health change programs. *Australian Review of Public Affairs*, 10(1), 23–40.

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, *7*(3), 93-99.

Fass, S., & Turner, K. (2015). The quantitative and qualitative content analysis of marketing literature for innovative information systems: the Aldrich Archive. *arXiv preprint arXiv:1505.04401*.

Feldman, R., & Sanger, J. (2006). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge, UK: Cambridge University Press.

Fox, S. (2011). *The social life of health information, 2011* (pp. 1-33). Washington, DC: Pew Internet & American Life Project.

Funk, T. (2011). *Social media playbook for business*. Santa Barbara, CA: Praeger.

Gao, Q., & Feng, C. (2016). Branding with social media: User gratifications, usage patterns, and brand message content strategies. *Computers in Human Behavior*, *63*, 868-890.

Gallaugher, J (2013). Information Systems: A Manager's Guide to Harnessing Technology, v. 1.4. 4th ed. London: Flat World Knowledge. 1-199.

Ghosh, S., & Desarkar, M. S. (2018). Class Specific TF-IDF Boosting for Short-text Classification.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, *21*(3), 267-297.

Groth, S. S., & Muntermann, J. (2011). An intraday market risk management approach based on textual analysis. *Decision Support Systems*, *50*(4), 680-691.

Gründemann, T., & Burghardt, D. (2016, April). Visual analysis of thematic, social and geospatial patterns of microblogging content using D3. In *LinkVGI workshop in association with the 19th AGILE conference on geographic information science*.

Hackler, D., & Saxton, G. D. (2007). The strategic use of information technology by nonprofit organisations: Increasing capacity and untapped potential. *Public Administration Review*, *67*(3), 474-487.

Hagenau, M., Liebmann, M., & Neumann, D. (2013). Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision Support Systems*, *55*(3), 685-697.

Hamilton, B. A. (2013). The field guide to data science. *Booze Allen Hamilton Web links*. Retrieved from https://www.boozallen.com/insight/2013/11/data-science-field-guide

Hambrick, M. E., Simmons, J. M., Greenhalgh, G. P., & Greenwell, T. C. (2010). Understanding professional athletes' use of Twitter: A content analysis of athlete tweets. *International Journal of Sport Communication*, *3*(4), 454-471.

Hanneman, R. A., & Riddle, M. (2005). Introduction to social network methods.

Harris, J. K., Mueller, N. L., Snider, D., & Haire-Joshu, D. (2013). Peer reviewed: Local health department use of Twitter to disseminate diabetes information, United States. *Preventing chronic disease*, *10*.

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464-472.

Heldman, A. B., Schindelar, J., & Weaver, J. B. (2013). Social media engagement and public health communication: implications for public health organisations being truly "social". *Public Health Reviews*, *35*(1), 13.

Heyer, G., Quasthoff, U., & Wittig, T. (2006). Text mining: Wissensrohstoff text. *W3l, Herdecke*, *18*.

Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management*, *6*(3-4), 248-260.

Jacobson, J., & Mascaro, C. (2016). Movember: Twitter conversations of a hairy social movement. *Social Media+ Society*, *2*(2), 2056305116637103.

Jha, A., Lin, L., & Savoia, E. (2016). The use of social media by state health departments in the US: analyzing health communication through Facebook. *Journal of community health*, *41*(1), 174-179.

Kent, M. L, & Taylor, M. (2002). Toward a dialogic theory of public relations. *Public Relations Review*, 28, 21-37.

Korda, H., & Itani, Z. (2013). Harnessing social media for health promotion and behavior change. *Health promotion practice*, *14*(1), 15-23.

Lacoste, S. (2016). Perspectives on social media ant its use by key account managers. *Industrial Marketing Management*, *54*, 33-43.

Leek, S., Canning, L., & Houghton, D. (2016). Revisiting the Task Media Fit Model in the era of Web 2.0: Twitter use and interaction in the healthcare sector. *Industrial Marketing Management*, *54*, 25-32.

Leek, S., Houghton, D., & Canning, L. (2017). Twitter and behavioral engagement in the healthcare sector: An examination of product and service companies. *Industrial Marketing Management*.

Li, C., & Bernoff, J. (2011). Groundswell: Winning in a World Transformed by Social Technologies. *Harvard Business Review Press*, 26(8), 286.

Lovejoy, K., & Saxton, G. D. (2012). Information, community, and action: How nonprofit organisations use social media. *Journal of Computer-Mediated Communication*, *17*(3), 337-353.

Lovejoy, K., Waters, R. D., & Saxton, G. D. (2012). Engaging stakeholders through Twitter: How nonprofit organisations are getting more out of 140 characters or less. *Public Relations Review*, *38*(2), 313-318.

Luo, B., Zeng, J., & Duan, J. (2016). Emotion space model for classifying opinions in stock message board. *Expert Systems with Applications*, *44*, 138-146.

Luo, J. S., & Smith, B. N. (2015). Social Media. In *Mental Health Practice in a Digital World* (pp. 183-197). Springer, Cham.

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, *23*(2), 254-277.

Mayring, P. 2000. Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research,* 1.

McCormick, T. Lee, H. Cesare, N. Shojaie, A. (2013). Using Twitter for Demographic and Social Science Research: Tools for Data Collection. *Center for Statistics and the Social Sciences*. 1 (1), 1-31.

Merchant, R. M., Elmer, S., & Lurie, N. (2011). Integrating social media into emergency-preparedness efforts. *New England Journal of Medicine*, *365*(4), 289-291.

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., & Hoving, C. (2013). A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research*, *15*(4).

Naaman, M., Boase, J., & Lai, C. H. (2010, February). Is it really about me?: message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 189-192). ACM.

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, *41*(16), 7653-7670.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science*, *31*(3), 521-543.

Neiger, B. L., Thackeray, R., Burton, S. H., Thackeray, C. R., & Reese, J. H. (2013a). Use of twitter among local health departments: an analysis of information sharing, engagement, and action. *Journal of medical Internet research*, *15*(8).

Neiger, B. L., Thackeray, R., Burton, S. H., Giraud-Carrier, C. G., & Fagen, M. C. (2013b). Evaluating social media's capacity to develop engaged audiences in health promotion settings: use of Twitter metrics as a case study. *Health promotion practice*, *14*(2), 157-162.

Niu, T., Zhu, S., Pang, L., & El Saddik, A. (2016, January). Sentiment analysis on multi-view social data. In *International Conference on Multimedia Modeling* (pp. 15-27). Springer, Cham.

Nkanunye, CC & Obiechina, GO (2017). Health Communication Strategies as Gateway to Effective Health Promotion and Well-being. *J Med Res Health Educ* 2017, Vol.1 No.3: 13.
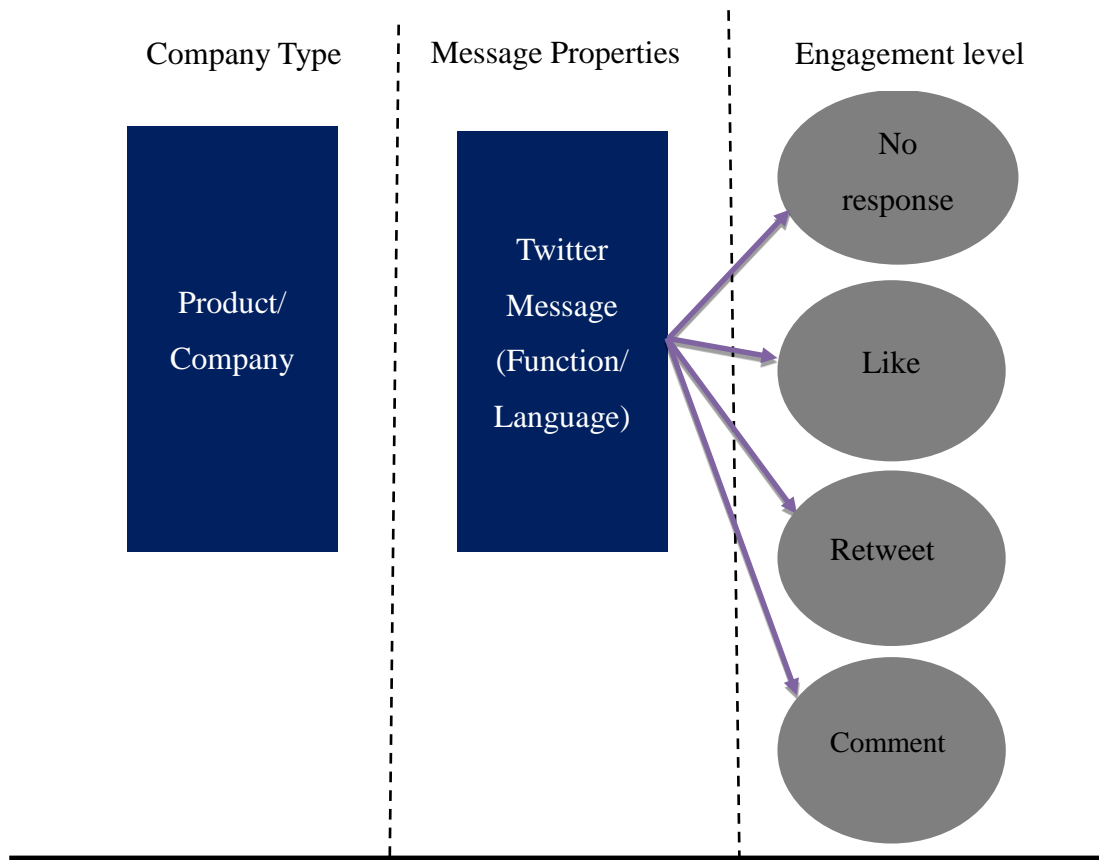
Osborne-Gowey, J. (2014). What is social media. *Fisheries*, *39*(2), 55-55.

Park, H., Rodgers, S., & Stemmle, J. (2013). Analyzing health organisations' use of Twitter for promoting health literacy. *Journal of health communication*, *18*(4), 410-425.

Park, H., Reber, B. H., & Chon, M. G. (2016). Tweeting as health communication: health organisations' use of twitter for health promotion and public engagement. *Journal of health communication*, *21*(2), 188-198.

Peramunetilleke, D., & Wong, R. K. (2002). Currency exchange rate forecasting from news headlines. *Australian Computer Science Communications*, *24*(2), 131-139.

Poushter, J; Bushop, C & Chwe, H. (2018) PEW. Social Media Use Continues to Rise in Developing Countries but Plateaus Across Developed Ones.

Priante, A., Need, A., Van Den Broek, T. A., & Hiemstra, D. (2018a). The role of Online Identity on Donations to Nonprofit Organizations in Online Health Campaigns. In *Academy of Management Proceedings*. Academy of Management Briarcliff Manor, NY 10510.

Priante, A., Ehrenhard, M. L., van den Broek, T. A., & Need, A. (2018b). Identity and collective action via computer-mediated communication: A review and agenda for future research. *New media & society*, *20*(7), 2647-2669.

Rabarison, K. M., Croston, M. A., Englar, N. K., Bish, C. L., Flynn, S. M., & Johnson, C. C. (2017). Measuring audience engagement for public health Twitter chats: insights from# LiveFitNOLA. *JMIR public health and surveillance*, *3*(2).

Ramos, J. (2003, December). Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning* (Vol. 242, pp. 133-142).

Rebholz-Schuhmann, D., Kirsch, H., Arregui, M., Gaudan, S., Riethoven, M., & Stoehr, P. (2007). EBIMed—text crunching to gather facts for proteins from Medline. *Bioinformatics*, *23*(2), e237-e244.

Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass communication & society*, *3*(1), 3-37.

Rus, H. M., & Cameron, L. D. (2016). Health communication in social media: message features predicting user engagement on diabetes-related Facebook pages. *Annals of Behavioral Medicine*, *50*(5), 678-689.

Rybalko, S., & Seltzer, T. (2010). Dialogic communication in 140 characters or less: How Fortune 500 companies engage stakeholders using Twitter. *Public relations review*, *36*(4), 336-341.

Safko, L. (2010). *The social media bible: tactics, tools, and strategies for business success*. John Wiley & Sons.

Saffer, A. J., Sommerfeldt, E. J., & Taylor, M. (2013). The effects of organisational Twitter interactivity on organisation–public relationships. *Public Relations Review*, *39*(3), 213-215.

Saxton, G. D., & Waters, R. D. (2014). What do stakeholders like on Facebook? Examining public reactions to nonprofit organizations' informational, promotional, and community-building messages. *Journal of Public Relations Research*, *26*(3), 280-299.

Scott, D. M. (2015). *The new rules of marketing and PR: How to use social media, online video, mobile applications, blogs, news releases, and viral marketing to reach buyers directly*. John Wiley & Sons.

Schultz, D.E. Peltier, J. (2013). Social media's slippery slope: challenges, opportunities and future research directions. *Journal of Research in Interactive Marketing*. 7(2) 86-99.

Sedley, R. (2008) Six theses on digital customer engagement in a troubled economy.

Shih, S. T., Carter, R., Heward, S., & Sinclair, C. (2017). Economic evaluation of future skin cancer prevention in Australia. *Preventive medicine*, *99*, 7-12.

Siamagka, N. T., Christodoulides, G., Michaelidou, N., & Valvi, A. (2015). Determinants of social media adoption by B2B organizations. *Industrial Marketing Management*, *51*, 89-99.

Smaldone, A., Stockwell, M.S., Osborne, J. C., Cortes, Y., Bekele, E., & Green, N.S. (2015).Adolescent and parent use of new technologies for health communication: A study in an urban Latino community. Journal of Public Health Research 2015; volume 4:376

Smith, A & Anderson, M (2018). A majority of Americans use Facebook and YouTube, but young adults are especially heavy users of Snapchat and Instagram.

Smith, B. G. (2010). Socially distributing public relations: Twitter, Haiti, and interactivity in social media. *Public relations review*, *36*(4), 329-335.

Sriram, B., Fuhry, D., Demir, E., Ferhatosmanoglu, H., & Demirbas, M. (2010, July). Short text classification in twitter to improve information filtering. In *Proceedings of the 33rd*

international ACM SIGIR conference on Research and development in information retrieval (pp. 841-842). ACM.

Statia (2017). Retrieved from https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of management information systems*, *29*(4), 217-248.

Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social media analytics. *Wirtschaftsinformatik*, *56*(2), 101-109.

Steinert-Threlkeld, Z. C. (2017). *Twitter as data*. Cambridge University Press.

Stempel, G. H, III. (2003). Content analysis, In G. H. Stempel III, D. H. Weaver, & G. C. Wilhoit (Eds.), *Mass communication research and theory* (pp. 209-219). Boston, MA: Pearson Education.

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010, August). Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Social computing (socialcom), 2010 ieee second international conference on* (pp. 177-184). IEEE.

Sullivan, S. J., Schneiders, A. G., Cheang, C. W., Kitto, E., Lee, H., Redhead, J., ... & McCrory, P. R. (2011). 'What's happening?'A content analysis of concussion-related traffic on Twitter. *British journal of sports medicine*, bjsports80341.

SunSmart (2018). Retrieved from http://www.sunsmart.com.au/

Swani, K., Brown, B. P., & Milne, G. R. (2014). Should tweets differ for B2B and B2C? An analysis of Fortune 500 companies' Twitter communications. *Industrial marketing management*, *43*(5), 873- 881.

Swani, K., Milne, G. R., Brown, B. P., Assaf, A. G., & Donthu, N. (2017). What messages to post? Evaluating the popularity of social media communications in business versus consumer markets. *Industrial Marketing Management*, *62*, 77-87.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24-54.

Thackeray, R., Neiger, B. L., & Keller, H. (2012). Integrating social media and social marketing: a four-step process. *Health promotion practice*, *13*(2), 165-168.

Thackeray, R., Neiger, B. L., Burton, S. H., & Thackeray, C. R. (2013a). Analysis of the purpose of state health departments' tweets: information sharing, engagement, and action. *Journal of medical Internet research*, *15*(11).

Thackeray, R., Burton, S. H., Giraud-Carrier, C., Rollins, S., & Draper, C. R. (2013b). Using Twitter for breast cancer prevention: an analysis of breast cancer awareness month. *BMC cancer*, *13*(1), 508.

Tsimonis, G., & Dimitriadis, S. (2014). Brand strategies in social media. *Marketing Intelligence & Planning*, *32*(3), 328-344.

Tuten, T. L., & Ashley, C. (2016). Do social advergames affect brand attitudes and advocacy?. *Journal of Marketing Communications*, *22*(3), 236-255.

Twitter. Tweet activity dashboard. URL: https://support.twitter.com/articles/20171990

Twitter (2016). "Selected Company Metrics and Financials." Technical Report.

Van den Heerik, R. A., Van Hooijdonk, C. M., Burgers, C., & Steen, G. J. (2017). "Smoking Is Sóóó... Sandals and White Socks": Co-creation of a Dutch anti-smoking campaign to change social norms. *Health communication*, *32*(5), 621-628.

Van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer engagement behavior: Theoretical foundations and research directions. *Journal of service research*, *13*(3), 253-266.

Waldherr, A., Heyer, G., Jähnichen, P., Niekler, A., & Wiedemann, G. (2016). 13 Mining Big Data With Computational Methods. *Political communication in the online world: Theoretical approaches and research designs*, 201.

Waters, R. D., & Jamal, J. Y. (2011). Tweet, tweet, tweet: A content analysis of nonprofit organisations' Twitter updates. *Public Relations Review*, *37*(3), 321-324.

Weber, R. P. (1990). Basic content analysis. Newbury Park, CA: Sage

Welch, V., Petkovic, J., Pardo, J. P., Rader, T., & Tugwell, P. (2016). Interactive social media interventions to promote health equity: an overview of reviews. *Health promotion and chronic disease prevention in Canada: research, policy and practice*, *36*(4), 63.

Wiersema, F. (2013). The B2B agenda: The current state of B2B marketing and a look ahead. *Industrial Marketing Management*, *4*(42), 470-488.

Wu, H., Chen, J., Zou, L., Zheng, L., Zhang, W., Meng, Z., ... & Sun, X. (2016). Community-based interventions to enhance knowledge, protective attitudes and behaviors towards canine rabies: results from a health communication intervention study in Guangxi, China. *BMC infectious diseases*, *16*(1), 701.

Xifra, J., & Grau, F. (2010). Nanoblogging PR: The discourse on public relations in

      Twitter. *Public Relations Review*, *36*(2), 171-174.

APPENDIX A: Leek *et al.*, (2017) Research scope

Figure 1. Research scope and focus of Leek, *et al*., 2017)

Company Type	Message Properties	Engagement level
Product/ Company	Twitter Message (Function/ Language)	No response / Like / Retweet / Comment

APPENDIX B: Codebook

Annotation Rules & Codebook

How to code

- Each Tweet can be assigned into *only one* of the defined categories for respectively.

- What to annotate: The description of tweets are in the "Text" and "Bio" columns.

- Annotate the number of perceived categorisation in the "Note 1" & "Note 2" column for Part I and Part II.

  o Part I: Identify user types as 1) Individual or 2) Organisation

  o Part II: Identify category of message function as one of : 1) Functional 2)Informational 3) Interactive 4) Promotional

PART I. TYPES OF USERS

Specific Rules

- What to annotate: The description of Tweet "Bio" column.

- The type of use for each tweet can be assigned into *only one* of the defined categories (i.e, either individual or organisation).

- Read "Bio" of each Tweet literally and follow the classification definition of types of users (Table 7) to annotate the Tweet category.

Table 7. Types of users classification

Category of users	Definition	Dimensions
1. Individual	The account is for personal use	
2. Organisation	The account is composed of individuals or represents not only one person, which can be recognised as one of the following entities: **institution; community association; foundation; for-profit corporation**	▪ Institution: Institution refers to an established order comprising rule-bound and standardized behaviour patterns (Jary and Jary 2000: 302), e.g., school. ▪ Community association: refers to a small or large unit formed through a group of people that have a common interest or aim. e.g., cancer association; sports clubs ▪ Foundation: a unit that gives or collects money to be used for special purposes, e.g., charity. ▪ For-profit corporation: a unit which aims to earn profit through its operations, e.g., online-shopping store.

PART II: MESSAGE FUNCTIONS

Specific Rules

- What to annotate: The description of Tweet "Text" column.
- Each Tweet can be assigned into *only one* of the defined categories.
- Read "Text" of each Tweet literally and follow the message function classification definition (Table 8) to annotate the Tweet category.
- In case a Tweet appears to serve dual purposes, codes are assigned to what is considered to be the *primary* purpose of the Tweet message.

Table 8. Message function classification

(a) Category of message function	Definition	(b) Sub-category	Examples
1. Functional	Message content contains **suggestions, advise, knowledge or solution** which might solve general problems or preventive approaches to a disease.	1.1 Useful information which might solve general problems: *E.g hats, sunscreen*	▪ Get #sunsmart with these tips advice #skinhealth always!! http://t.co/P71vmML3M5' ▪ #SunSmart tips: wear sunglasses a broad-brimmed hat when outside to shade your face, neck, ears from harmful UV.
		1.2 Suggestions or advise: *E.g make sure, remember, don't*	▪ It's a sunny day again here in Reading, don't forget to be #SunSmart and pop sun cream on like Palmer! #ReadingRelay http://t.co/f3H4ZIHRBq
		1.3 Disease-associated: *E.g melbourne, skincancer*	▪ You are at higher risk of melanoma if you Have red or blonde hair. #SPF be #SunSmart
2. Informational	Message content is related to share information, which can be about **an upcoming or past event, personal experience,** or simply an **opinion** or **preference**.	2.1 An upcoming or past event *E.g im703busso, event, triathlon,*	▪ Sunsmart Kids Triathlon 2014 #sunsmart #im703busso http://t.co/p1NTZmYJm5
		2.2 Personal experience	▪ My sister just sent me photos of her sun burn!! She should know better - I ve been going on about being #sunsmart for years! #NotImpressed
		2.3 An opinion or preference *E.g great*	▪ These 2 great products will give you peace of mind that your baby is happy &amp safe in the sun http://t.co/lTrstuehcg #baby #sunsmart'

3. Interactive	Message contains a unique **interactive symbol "@"**, or sentences with **open questions** which lead people to discuss or **thanking**.	3.1 Mention a certain account user with the simple "@" in text body *E.g: @*	▪ @RealHughJackman Slip Slop Slap Seek and Slide #sunsmart
		3.2 Open-question *E.g are you…?, should we?, do you?, can you?*	▪ Do you have regular skin screening tests? #SunSmart #Melanoma
		3.3 Thanking *E.g thank you*	▪ We love all of our fans! Thank you for being #sunsmart!
4. Promotional	Message shows **strong intentions or physical incentives** that lead people to the good causes of **participating** services or activities.	4.1 Participation: encourages people to take part in a certain activity, contest or sign up a team. *E.g join, "save your skin, save your life", check out*	▪ Join our mailing list for fashion, lifestyle and daily shade inspiration http://t.co/Q5mU7lHouz #sunsmart #melanoma http://t.co/EhXUDV3Y2v ▪ Checkout the action from the @BTI_Aus #WA Junior Surfing Titles presented by #SunSmart at Trigg Beach https://t.co/IfTYRh8Ndb #wasurfers
		4.2 Incentive: encourages people to buy a certain product by monetary support, giveaway or coupon. *E.g free, giveaway*	▪ Really wish that #TheWarehouse can hurry up and do a buy 1 get 1 free sale on sunscreens!! #goingred #burnt #soexpensive #sunsmart'

APPENDIX C: Keywords & Thematic topics among message functions

Distinctive keywords and thematic topics are shown per different message functions

To have an overall insight into the communication during the SunSmart, we analysed original tweets and Retweets (n=2381) to know what keywords and thematic topics constitute each category of message function. The results in Table 9 indicate that the keywords in different message functions are much varied. For instance, "stay", "index", "protected" and "common" are the most important keywords in the *Functional* message, whereas these words are relatively insignificant in regard to other message types. The main themes of the four message functions are interpreted from the hashtags visualisation analysis chart in Appendix E-H.

Table 9. Summary of keywords and thematic topics among message functions (*Original tweets & Retweets*)

Message function	Keywords	Main thematic topics
Method	TF-IDF via Natural Language Processing	Hashtag Visualisation Thematic Analysis
Functional	. stay . index . protected . common	Theme 1: *general skin health issues and skin cancer* Theme 2: *sunscreen usage* to *avoid hydration in LA* Theme 3: *wintertime sun protection* Theme 4: *sun protection in Melbourne*
Interactive	. easy . stylish . enter . win	Theme 1: *activity, sun protection* Theme 2: *health disease prevention* Theme 3: *skin cancer* Theme 4: *product marketing* Theme 5: *community marketing*
Informational	. palmtrees . sunbeds . maldives . nofilter	Theme 1: *event-related* Theme 2: *sun protection in vacation-physical block.* Theme 3: *sun protection in vacation-chemical block.*
Promotional	. slip . free . download . shipping	Theme 1: *inspiring slogan* Theme 2: *bamboo product promotion* Theme 3: *sunscreenband promotion* Theme 4*: a football activity in Montclair*

APPENDIX D: Keywords of message functions in Original tweets

Table 10 illustrates a comparison of original tweets by organisations and individuals. The results indicate that both types of users post *Interactive* messages with keywords "stylish" and "enter". However, with those two terms being the exceptions, individuals and organisations used different keywords in their posted messages per different message functions.

Table 10. Keywords of message functions in *Original tweet*－individuals vs. organisations

Message Function	Original Tweet	
	Individuals	Organisations
Functional	. index . 11 . listentoyourmother . lobstered	. uvindex . dont . wear . remember
Interactive	. easy . stylish . enter . win	.sunscreenbands . stylish . enter . thanks
Informational	. margsy . springbreak . holiday . nofilter	. 2014 . schoolsurfingtitles . triathlon . im703busso
Promotional	. slop . slip . pldtsme . slipslopslap	. join . download . free . shipping

*Note: the difference between individuals and organisations are highlighted in *red*

APPENDIX E: Functional: Top 150 nodes

Figure 8. *Functional*: Top 150 nodes
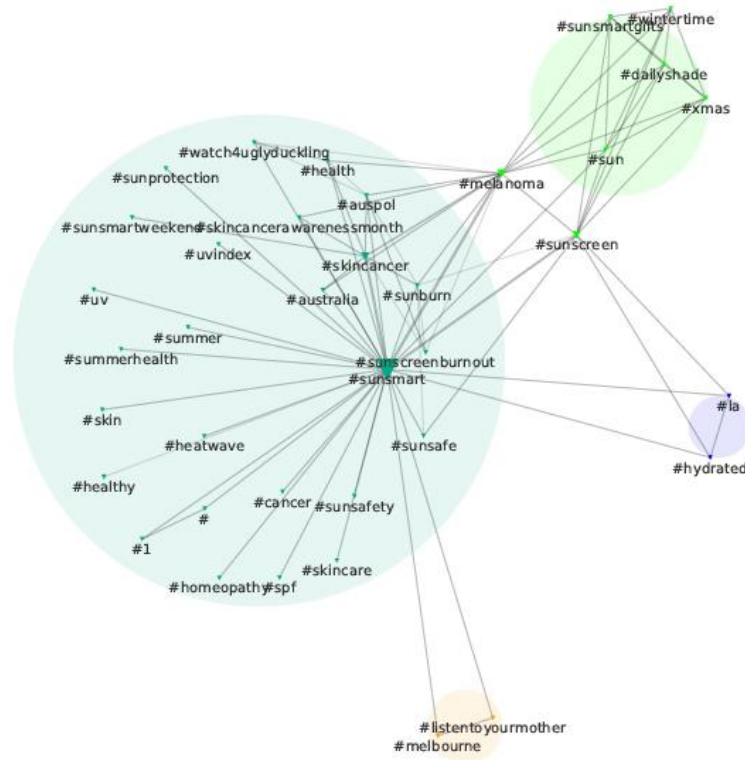
Primary themes among the messages in the Functional function (Original tweets & Retweets)

Figure 8 shows the three most weighted nodes in the *functional* message category are (i.e. #sunscreen #melanoma, and #skincance, and four distinct clusters contain:

- Cluster 1 (blue) deals with issues regarding *general skin health issues and skin cancer*, which is composed of various single hashtags mentioned along with #sunsmart, such as #sunscreenburnout, #sunsafe, and #sunscreen.
- Cluster 2 (purple) presents a subject tending to be *sunscreen usage t*o *avoid hydration in LA*, and related hashtags such as #hydrated, #la and #sunscreen.
- Cluster 3 (green) refers to *wintertime sun protection* as some related wording is used, such as #xmas, #wintertime, and #wintertimegifts along with #melanoma and #sunscreen.
- Cluster 4 (orange) is related to *sun protection in Melbourne*, and this topic group is in a remote position on the map.   .

Thematic topics among the messages in the Functional message type: Original tweets

Figure 9 illustrates that "*general skin health issues, and skin cancer*" (theme 1) is the principal theme for both types of user. However, they focus on different small topics (hashtags). In addition, the "*sun-protection in Melbourne*" (theme 4) is the second most important theme for individuals to frame original tweets, whereas "*wintertime sun-protection*" (theme 3) is more critical to organisations.

Figure 9. Thematic topics in the *Functional* message function of *Original* tweets

Individual	Organisation
Theme 1: general skin health issues and skin cancer. ⇨ #melanoma, #heatwave; #skincancer. *Theme 4: sun-protection in Melbourne* ⇨ #listentoyourmother #sunscreen, #melbourne	*Theme 1: general skin health issues and skin cancer* ⇨ #sunsafe, #uvindex, #summerhealth, #sunscreen, #skincancer *Theme 3: wintertime sun-protection* ⇨ #wintertime, #xmas, #sunsmartgifts, dailyshade

APPENDIX F: Interactive: Top 150 nodes

Figure 10. *Interactive*: Top 150 nodes
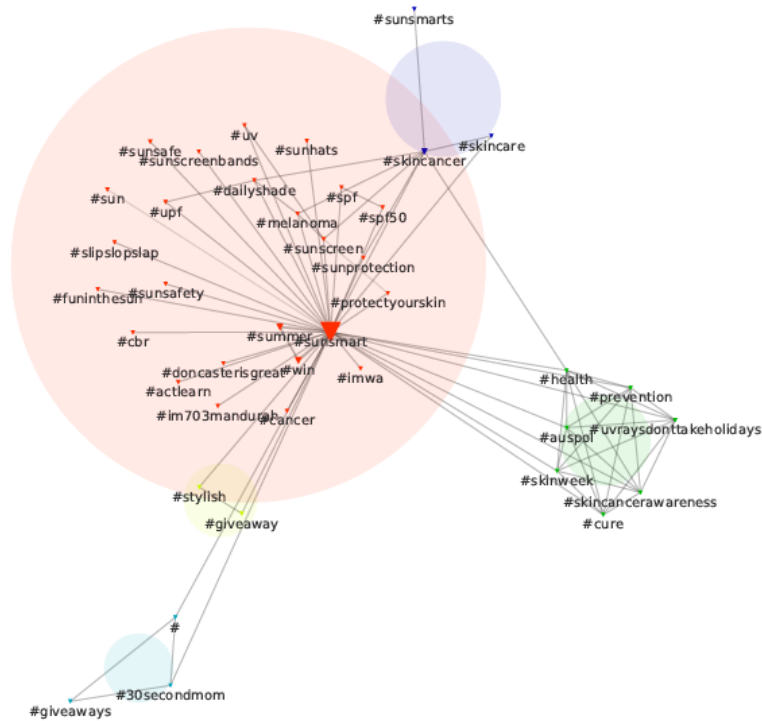
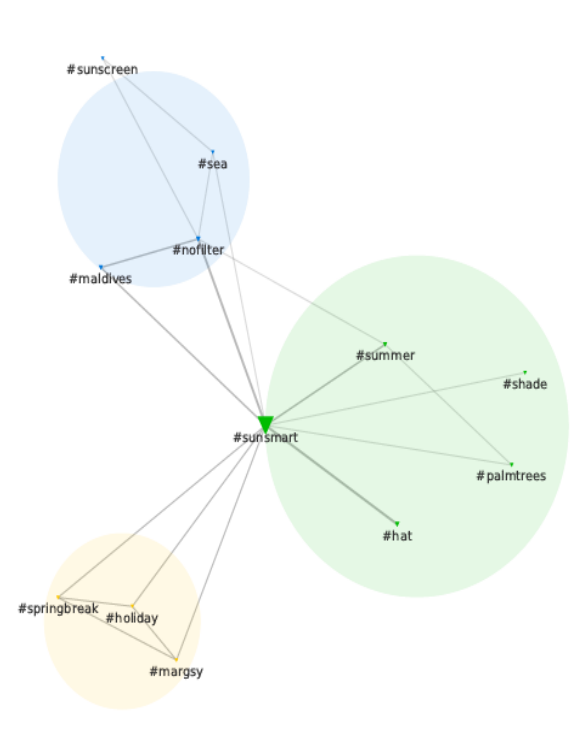*Primary themes among the messages in the Interactive function (Original tweets & Retweets)*

Figure 10 shows the four most weighted nodes (i.e. #win, #summer, #skincancer, and #skincare) in *interactive* message and five distinctive clusters:

- Cluster 1 (red) is a broad theme that incorporates *activity, sun protection*; activity constituted mainly by two essential terms #win and #summer as well as several single hashtags, such as #imwa, #im703man.

- Cluster 2 (green) presents an interweaved nest conveying the *health disease prevention* theme with a collection of common hashtags- #health, #cure, #skincancerawareness, #uvraydonttakeholldays.

- Cluster 3 (blue) conveys a *skin cancer* theme, composed of the #skincancer and #skincare hashtags.

- Cluster 4 (yellow) is related to *product marketing*, containing #stylish and #givewaway.

- Cluster 5 (light blue) is interpreted as *community marketing*, including the terms #giveaway and #30secondmom.

Thematic topics in the Interactive message type: Original tweets

Figure 11 indicates that although both individuals and organisations discussed theme 1 and theme 4, the topics (i.e. hashtags) of composition are different. Individuals used a great deal of #win and #summer to frame the theme 1, whereas organisations less framed their posts by these two terms in an interactive message. In addition, while both parties deliver *interactive* messages with *product marketing* (theme 4), their proximity reveals that this subject is more crucial to individuals than organisations.

Figure 11. Thematic topics in the *Interactive* message function of *Original tweets*

Individual	Organisation
Theme 1: *activity, sun protection and skin cancer* ⇨ #summer, #win, #melanoma *Theme 3: skincancer* ⇨ #skincancer, #health, *Theme 4: product marketing* ⇨ #stylish, #giveaway	*Theme 1: activity, sun protection and skin cancer* ⇨ #spf 50, #spf, #sunhats, #win, #summer *Theme 2: health-disease prevention* ⇨ #sunprotection, #dailyshade *Theme 4: product marketing* ⇨ #stylish, #giveaway

APPENDIX G: Informational: Top 100 nodes

Figure 12. *Informational*: Top 100 edges
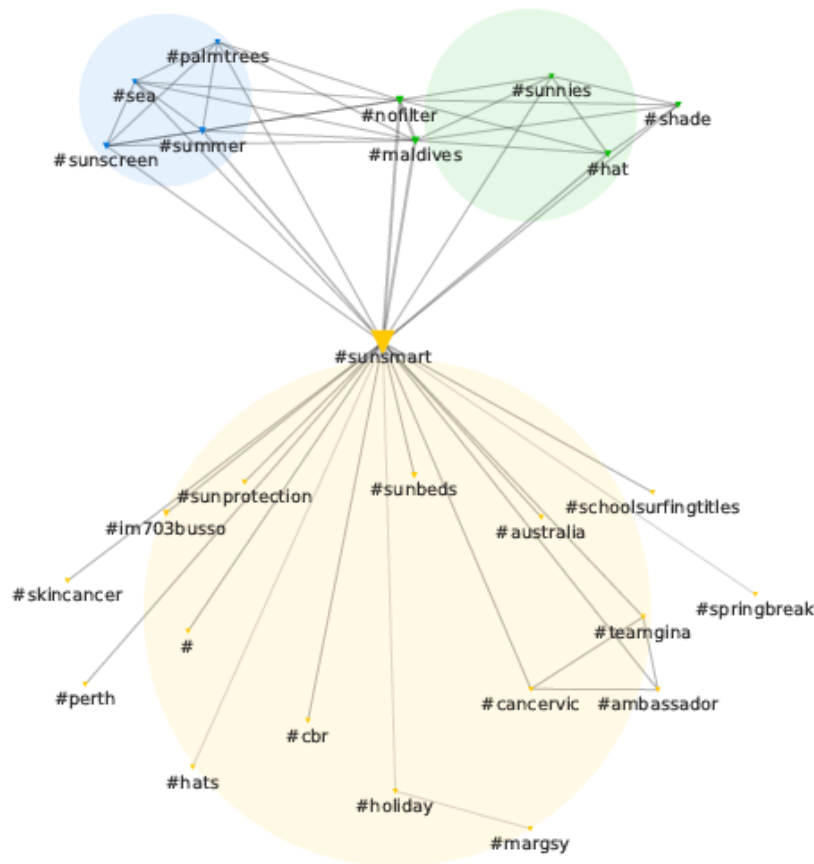
Primary themes among the messages in the Informational function (Original tweets & Retweets)
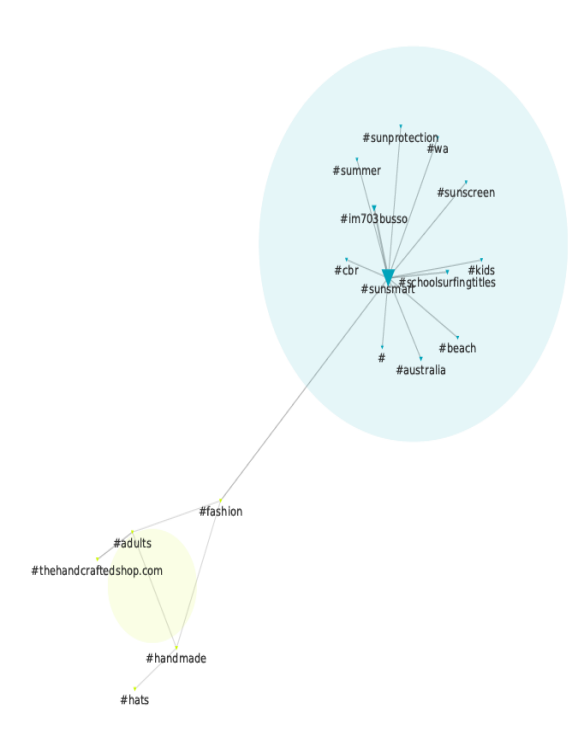
Figure 12 illustrates the four weighted nodes (i.e. #nofilter, #maldives, #summer, and #hat) in *Informational* message type and three distinct clusters:

- Cluster 1(yellow). This primary cluster seems to be *event-related,* including terms such as #im703busso, #holiday, #cancervic, #springbreak.
- Cluster 2 (green) and cluster 3 (blue) reveal a similar theme *sun-protection on vacation.* The difference is, theme 2 tended to be mentioned as *physical-block,* such as #shade and #hats, and theme 3 highlighted the *chemical-block* (i.e. #sunscreen).

Thematic topics among the messages in the Informational message type: Original tweets

Figure 13 shows that both parties mostly posted information about theme 1 and theme 3; however, their focuses still differ. For *event-related* information (theme 1), personal users preferred to publish holiday information (i.e., #spring break, holiday), but organisations disseminate information about sports campaigns "#im703busso". While theme 3 was posted by both parties, hashtag #sunscreen is less important in personal tweets than in organisational tweets.

Figure 13. Thematic topics in the *Informational* message function of *Original tweets*

Individuals	Organisations
Theme 1: event-related ⇨ #springbreak, #holiday, #margsy	*Theme 1: event-related* ⇨ #im703busso
Theme 2: sun protection on vacation-physical block. ⇨ #shade, #hat, #summer	*Theme 3: sun protection on vacation-chemical block.* ⇨ #sunscreen, #sunmmer, #sunprotection
Theme 3: sun protection on vacation-chemical block. ⇨ #maldives, #sunscreen, #nofilter	

APPENDIX H: Promotional: Top 50 edges

Figure 14. *Promotional*: Top 50 edges

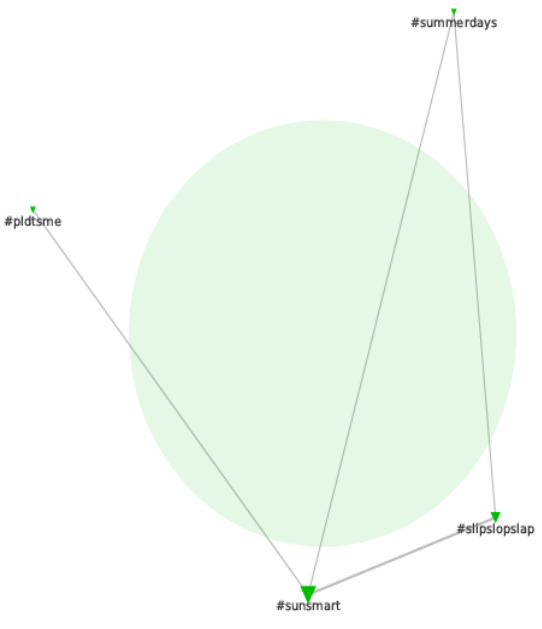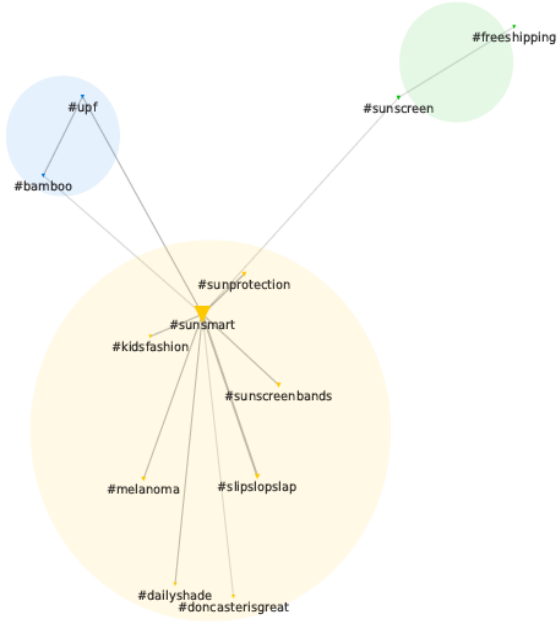***Primary themes among the messages in the Promotional function (Original tweets & Retweets)***

Figure 14 illustrates that there are two weighted hashtags (i.e. #slipslopslap, #doncasterisgreat) and four major themes in the *Promotional* message category:

- Cluster 1 (orange) is concerned with the *inspiring slogan,* which is mainly constituted by the #slipslapslop and #doncasterisgreat hashtags.
- Cluster 2 (green) seems to be the *promotion for sun-protection products made of bamboo,* containing hashtags #bamboo, #upf, #melanoma.
- Cluster 3 (blue) is related to *sunscreenband product promotion*, containing hashtags #30secondmom, #sunscreeband, #giveaways.
- Cluster 4 (light blue) is *a football activity in Montclair* which is a township in New Jersey, and this theme was composed of #mhtchairfootball and #mhsmounties.

Thematic topics among the messages in the Promotional message type: Original tweets

Figure15 shows that organisations used a wide range of themes to construct promotional posts, whereas individuals mainly framed posts via *inspirational slogan* (theme 1)

Figure 15. Thematic topics in the *Promotional* message function of *Original tweets*

Individuals	**Organisations**
Theme 1: *inspiring slogan* ⇨ #slipslapslop, #summerdays	*Theme 1: inspiring slogan* ⇨ #slipslapslop *Theme 2*: *bamboo product promotion* ⇨ #bamboo, #upf *Theme 3: sunscreenband promotion* ⇨ #sunscreeband *Other theme: sunscreen product promotion* ⇨ #sunscreen, #freeshipping