

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of
Business

Lee Kong Chian School of Business

12-2020

Social media influencers and Instagram storytelling: Case study of Singapore Instagram influencers

Mark CHONG

Singapore Management University, markchong@smu.edu.sg

GOTTIPATI Swapna

Singapore Management University, SWAPNAG@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research



Part of the [Business and Corporate Communications Commons](#), [Communication Technology and New Media Commons](#), and the [Social Media Commons](#)

Citation

CHONG, Mark and GOTTIPATI Swapna. Social media influencers and Instagram storytelling: Case study of Singapore Instagram influencers. (2020). *Journal of Applied Business and Economics*. 22, (10), 81-96. Available at: https://ink.library.smu.edu.sg/lkcsb_research/6828

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Social Media Influencers and Instagram Storytelling: Case Study of Singapore Instagram Influencers

Mark Chong
Singapore Management University

Swapna Gottipati
Singapore Management University

While the use of social media influencers (SMIs) by brands is becoming more widespread, the academic literature about SMI communication is still scarce. This is one of the first studies on SMI brand storytelling, using data mining and natural language processing to understand how SMIs tell brand stories on Instagram, what kinds of stories they tell, and the impact they have on follower engagement. The findings show that the “rise-fall” emotional arc was the most common story arc used by SMIs. In addition, SMIs frequently used the first-person perspective and featured themselves as the protagonists in their stories. Last, SMIs who used more first-person pronouns and more positive emotions in their stories received more “likes” than those who used fewer first-person pronouns and fewer positive emotions. The paper concludes with a discussion of the study’s implications for SMI communication theory-building and practice as well as its limitations.

Keywords: social media influencers, narrative persuasion, storytelling, Instagram, data mining

INTRODUCTION

As consumers increasingly turn to peer-to-peer recommendations and communication (Hughes, Swaminathan, & Brooks, 2019), organizations are increasingly integrating social media influencers (SMIs) into their marketing strategies to achieve business objectives (Bakker, 2018; de Veirman, Cauberghe, & Hudders, 2017). SMIs have been defined as “people who have built a sizeable network of people following them” (de Veirman et al., 2017, p. 798). In addition, they are regarded as “a trusted tastemaker in one or several niches (de Veirman et al., 2017, p. 798). Unlike mainstream celebrities, influencers are perceived to be more accessible, believable, and thus easy to relate to — they share insights into their personal, everyday lives, experiences and opinions with their followers through blogging, vlogging and short-form content (Abidin, 2016; Schau & Gilly, 2003).

SMIs are highly valuable to organisations for a number of reasons. First, they have access to younger audiences through diverse social media platforms and to customers who are more difficult to reach through traditional media (Chatzigeorgiou, 2017; Lin, Bruning, & Swarna, 2018; Carter, 2016; Freberg, Graham, McGaughey, & Freberg, 2011). Second, SMIs’ endorsements are highly personal and interwoven into the textual and visual narratives of their personal lives, so they will likely be perceived as unbiased

opinions (Abidin, 2015) — their higher perceived authenticity and credibility may lead to lower resistance to their messages (de Vries, Gensler & Leelang, 2012).

The goals of SMI communication can range from increasing brand awareness and improving brand recognition to generating social buzz around a brand in order to boost brand value and revenues (Sudha & Sheena, 2017). Sponsored blogging, in which companies solicit influencers to post about specific products or brands, is already a large and important category of influencer marketing (Hughes et al., 2019). Spending on influencer marketing is projected to reach \$10 billion by 2020, with 65% of multinational brands planning to increase their spending on influencer marketing (Belton, 2019; Mediakix, 2018). However, the effectiveness of SMIs is still low – the average engagement rate per post for an influencer on Facebook is 0.37% and 0.05% on Twitter (Feehan, 2018).

The rise of SMIs has been aided in part by the rapid growth of Instagram and its user base in recent years (Marwick, 2015; Senft, 2013). SMIs typically start out as “ordinary” consumers who then move on to become opinion leaders (Song, Cho, & Kim, 2017) and “third-party endorsers” (Freberg, et al., 2011, p. 90) through self-branding strategies (Erz & Christensen, 2018) that involve reaching out to followers and disclosing their personal information (Marwick, 2015). As SMIs are perceived to be more credible and relatable than conventional celebrities, they are often engaged by organizations to promote brands, products and services on Instagram (Djafarova & Rushworth, 2017), especially as they can influence both pre- and post-purchase decisions of consumers (e.g. Song et al., 2017).

While the use of SMIs by brands is becoming more widespread (Backaler, 2018), the academic literature about SMI communication is still scarce and rather fragmented (Bakker, 2018; Voorveld, 2019). Furthermore, even though advertising, branding and communication research has shown how companies can benefit by telling stories to customers (van Laer, De Ruyter, Visconti, & Wetzels, 2014), research on SMI storytelling is scarce. Moreover, although the role of narrative transportation in advertising persuasiveness has been well established (van Laer et al. 2014), there have been very few studies to date on how the theory can advance our understanding of SMI communication on social media.

This study is a response to the call for more research into SMI communication, including applying a broader range of empirical methods (Sundermann & Rabe, 2019). This is possibly the first study on SMI brand storytelling, using data mining and natural language processing to understand: 1) How SMIs tell stories on Instagram; 2) What kinds of stories SMIs tell; and 3) The relationship between SMI storytelling and follower engagement. The results of this study have both theoretical and practical implications for SMI communication theory-building and practice.

LITERATURE REVIEW

Narrative Transportation

Brand stories as a promotional platform is rising in importance to both consumers and firms (Hamby, Brinberg, & Daniloski, 2019). Much of this interest has undoubtedly been driven by recent findings in neuroscience and cognitive science (e.g. Zak, 2015; Kahneman, 2013; Heath & Heath, 2007) that link stories and storytelling to various persuasive outcomes. People process narratives less critically (Adaval & Wyer, 1998; Escalas, 2007). In contrast, they engage in critical scrutiny and argument-focused elaboration when they process argument-based messages (Petty & Cacioppo, 1986).

Narrative transportation theory (Green & Brock, 2000) offers a useful theoretical lens for understanding brand storytelling by SMIs. Narrative transportation has found many applications in advertising research (for review, see van Laer et al., 2014) and been used to explain the persuasive effects of (advertising) stories on consumers (Gerrig, 1999; Green, 2008; Wang & Calder, 2006, 2009). The term “narrative transportation refers to the audience's pleasurable and engrossing experience of being swept away by a story (Gerrig 1999; Green & Brock, 2000; van Laer, et al. 2014) such that “the world of origin [is] partially inaccessible to the story receiver, thus marking a clear separation in terms of here/there and now/before, or narrative world/world of origin” (van Laer et al., 2014, p. 799). This experience is the consequence of narratives focusing audiences' attention on the story world while simultaneously reducing their processing of their immediate environment (Gerrig, 1999; Green & Brock, 2000). Busselle and Bilandzic (2008)

suggest that the experience of feeling transported into - and getting lost in - a narrative involves audience members substituting their sense of time and place from the real world with the time and place represented in the story world.

Scholars have shown that a story can engross the story receiver in a transformational experience (Phillips & McQuarrie, 2010) that results in narrative persuasion, which manifests itself in changes in story receivers' affective and cognitive responses, beliefs, attitudes, and intentions (van Laer et al., 2014). More specifically, narrative persuasion has been associated positive attitudes toward the story, story-consistent affective responses, story-consistent attitudes toward the experience described in the narrative, and story-consistent beliefs, attitudes and intentions (Argo, Zhu, & Dahl 2008; Escalas, 2007; Wang & Calder 2006, 2009; van Laer et al., 2014). In addition, narrative transportation may make consumers perceive an advertising story to be more truthful (Green & Donahue, 2011). Greater transportation has also been systematically associated with a more positive evaluation of a story's protagonist (Green & Brock, 2000). As liked sources can be more effective in inducing attitude or belief change (Eagly & Chaiken, 1993), transported individuals can feel greater affinity for story protagonists and thus be more vulnerable to being influenced by the feelings or beliefs expressed by those characters (Green & Brock, 2000). Transportation could theoretically be elicited by any text. Nonetheless, the evidence indicates that the likelihood of experiencing it is higher in response to narratives, especially well-crafted ones (Green & Brock, 2000).

While there are numerous definitions of story, this study restricts story to mean "a storyteller's account of an event or a sequence of events leading to transition from an initial state to a later state or outcome" (Van Laer et al., 2014, p. 798). According to this account, a story would comprises the following elements: (1) a plot that frames the temporal sequence of events; (2) the characters that play a role in the plot; (3) the resolution of the plot, which is typically derived from the character's resolution of a conflict or tension (Banerjee & Green, 2012; Green & Brock, 2000; Phillips & McQuarrie, 2010; Stern, 1994). A story plot influences narrative transportation by provoking mental imagery (Escalas, 2004a; Green & Brock, 2000) – it is through a mentally imagined plot that a story resembles real-life experiences (Green, 2006).

Story plots often follow distinct emotional trajectories or arcs, which trace the ups and downs of a protagonist's/character's journey (Reagan, Mitchell, Kiley, Danforth, & Dodds, 2016). A big-data analysis of Project Gutenberg's fiction collection shows that the most common emotional arcs are 'rags to riches' (rise), 'tragedy' (fall), 'man in a hole' (fall-rise), 'Icarus' (rise-fall), 'Cinderella' (rise-fall-rise), and 'Oedipus' (fall-rise-fall) (Reagan et al., 2016). Notably, the same study discovered that two out of the three emotional arcs that are most popular with readers – i.e. 'Cinderella', 'Oedipus' – feature the most change in direction (Reagan et al., 2016). Similarly, van Laer et al. (2019) found that the two story genres which have the most change in their arcs – i.e. comedies and tragedies – transport and persuade consumers more than other. Zak (2015) has shown that stories with a 'dramatic' arc, as compared with 'flatter' stories, increase oxytocin and cortisol levels in the audience's bloodstream. This leads to the study's first research question:

RQ1: What are the most common emotional arcs?

RQ2: Which emotional arcs trigger the highest audience engagement (measured by the number of likes)?

Brand Management

Identifiable characters are another antecedent of narrative transportation (van Laer et al., 2014). According to Hoeken, Kolthoff and Sanders (2016), the persuasive effects of story emanate from identification, a process in which an audience assumes the perspective of the main character in a story and perceives the story from that character's vantage point (Busselle & Bilandzic, 2008). Identifiable characters affect narrative transportation because story receivers vicariously experience characters' thoughts, feelings, beliefs, attitudes, and actions and empathize with them (van Laer et al., 2014; Slater & Rouner 2002). Escalas, Moore, and Britton (2004) show that identifiable characters make story receivers more likely to be transported into commercial stories. Identification with story characters has also been shown to foster self-brand connections (Escalas, 2004b).

Sachs (2012) posited that brands (and by extension, influencers) should play the role of mentors in their consumers' transformation journeys, supporting and guiding them through the challenges they face and contributing to the fulfilment of their individual goals. In narratology, the hero's journey is the common template of a broad category of tales that involves a hero who goes on an adventure, and in a decisive crisis wins a victory, and then comes home changed or transformed. In this story framework, the protagonist in the story is helped by a mentor (Campbell, 1949). Accordingly, the focus in brand stories should shift from the brand's powers and solutions to the customers' efficacy. In this model, audiences and customers become the central characters — i.e. the heroes — in the brand stories while brands (and by extension, influencers) become the enablers and supporters (Adi, Crisan, & Dinca, 2015). To date, there has been no empirical study on the actual prevalence of this practice in brand communication.

Hashtags may provide insights into the narrative role played by SMIs in their own Instagram stories. First developed as a user innovation on Twitter, a hashtag does not only indicate or highlight the topic of a post (i.e. what the post is “about”) – it can also allow the poster to communicate implicit or explicit meanings to the audience (Scott, 2015, 2018; Erz & Christensen, 2018). In this way, a poster can use “the tag space to guide the reader towards the intended interpretation” (Scott, 2018, p. 3). This discussion forms the basis for the first hypothesis:

H1: SMIs more frequently position their audiences (rather than themselves or the brands they champion) as the protagonists in their stories.

Brand storytellers can induce identification by using narrative perspective to make audiences ‘see’ the story through the eyes of the protagonist (or take a spectator’s view). The choice of personal pronouns is a well-established way to guide this form of cognitive perspective taking (Sanford & Emmott, 2012; Bergen & Chang, 2005; Brunye, Ditman, Mahoney, Augustyn, & Taylor, 2009; Brunye, Ditman, Mahoney, & Taylor, 2011; Ditman, Brunye, Mahoney, & Taylor, 2010). First-person pronouns are known to facilitate a more immediate experience and thus stronger identification (Oatley, 1999; Papeo, Corradi-dell’acqua, & Rumiati, 2011). Furthermore, first-person pronoun stories are more likely to induce narrative transportation than third-person pronoun stories (Hartung, Burke, Hagoort, & Willems, 2016). This leads to the following hypothesis:

H2: SMIs who use more first-person pronouns will receive more “likes” than SMIs who use fewer first-person pronouns.

Hedonic Content and Audience Engagement

Research shows that hedonic content, which refers to the enjoyment and emotions that an individual experiences from reading or viewing a piece of content (Hughes et al., 2019), can influence consumer attitudes and word of mouth (Berger & Schwartz, 2011; Kim, Ratneshwar, & Thorson, 2017; Mikyoung et al., 2019). Specific emotions can trigger arousal, which in turn boosts the virality of online content (Berger & Milkman, 2012). Building on these findings, Hughes et al. (2019) showed that content featuring high hedonic value is positively related to the number of blog post comments and the number of Facebook likes. Using Hollebeek’s (2011) definition of engagement as a “customer’s cognitive, emotional, and behavioural activities” (p. 555), this study focused on user interactions with SMI content, which are operationalized as ‘likes’ on SMI’s Instagram posts. This leads to our third hypothesis:

H3: SMIs who use more positive emotions in their narratives will receive more “likes” than SMIs who use more negative emotions.

METHOD AND DATA

In this section, we first provide the background of the text analytics API, LIWC and then describe the details of the data. Finally, we describe the solution methodology we developed to address the hypothesis testing and answering research questions.

LIWC Background

Linguistic Inquiry and Word Count (LIWC) is a computerized text analysis software program and is an appropriate tool for studying the word constructions of an Instagram post. This program calculates the percentages of words and terms in a text that fall into certain categories (Pennebaker & Graybeal, 2001). According to Tausczik and Pennebaker (2010), natural language use reveals people's disposition regarding their processing and interpretation of information to make sense of the environment.

LIWC involves a word count approach, calculating word frequencies and converting them into percentages for various categories. Its dictionary covers about 80% of the vocabulary that people use in everyday written and spoken language (Pennebaker & Graybeal 2001). The LIWC 2007 word categories were created with the goal of delineating useful classes of words for the description of formal aspects of language use (e.g., use of pronouns; past, present, future tense; use of punctuation), its functional aspects (e.g., words signifying causation, certainty, negation, temporal and spatial relations), reference to emotion (e.g., words signifying positivity, anxiety, anger), reference to fundamental perceptions and actions (e.g., words related to the senses) and reference to social categories (family, friends), references to biological aspects (health, body, sexuality, food) and concerns (e.g., words related to leisure, work, money, home). Its average classification rate of any given words in various writing and speech texts is 86% (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). For most of the categories, mean values indicate the mean percentages of all of the words that participants used that fell into a particular category. For example, a mean score of 5.43 for positive emotion words indicates that 5.43% of the words participants used were associated with positive emotion (e.g., love, nice, sweet).

According to the design of our study, four categories and their subcategories were coded and computed. These categories comprised pronouns (e.g., I, we, mine), social (e.g., daughter, buddy; subcategories: family, friend), affective (e.g., happy; subcategories: positive and negative emotions, which comprise anxiety, anger, and sadness), biological (e.g., eat, pain, dish; subcategories: body, health, ingestion) and personal concerns (e.g., job, cook, kitchen, church; subcategories: work, leisure, home, religion).

Data

The latest 1000 posts of the top 22 Instagram influencers in Singapore (Lim, 2019) were analyzed in our models – two of the original list of 24 influencers were excluded, as they had fewer than 10,000 followers. In total, 20318 posts were collected. It should be noted that some influencers had fewer than 1000 posts. If a post was shorter than three sentences, it was filtered out from the study.

Furthermore, if the last sentence was made up of only a hashtag, it was also filtered out from the data. This is because a hashtag appearing as the very last segment of a post makes it very difficult to ascertain the emotion of the post (if the post comprises fewer than five sentences). A total of 7369 posts were left after this process of filtering. The statistics of the influencers' information for this study are shown in Table 1.

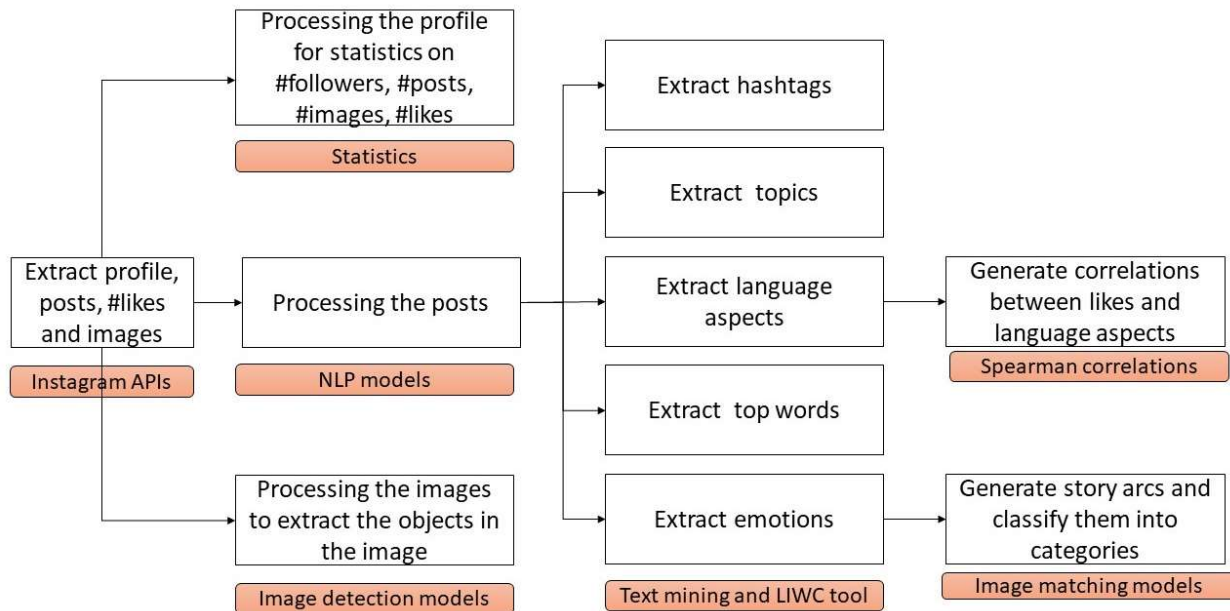
TABLE 1
DATA STATISTICS USED IN THE STORY

Category	#Influencers	Avg #of Followers	Total # of posts	Total # of posts collected (before filtering)	Total # of posts collected (after filtering)
Beauty	3	178.3k	19566	3000	754
Fashion	3	200.4k	9811	3000	828
Fitness	3	46.0k	2557	2270	668
Food	3	194.7k	7176	2569	1918
Lifestyle	3	321.3k	5688	2677	841
Parenting	3	12.8k	9892	3000	1282
Pets	1	20.0k	802	802	365
Travel	3	111.0k	6956	3000	713

Method

Figure 1 depicts the detailed solution model to extract, analyse and gain insights of posts from Instagram influencers. The inputs to the model are the Instagram content extracted using the APIs from Instagram. At every stage, the diagram depicts the task (blue) and the techniques (orange) used for mining the data.

FIGURE 1
SOLUTION MODEL



Instagram APIs

Data for this study was collected from the public Instagram API <https://www.instagram.com/developer>. A python code was created to access the SMI profiles, posts and images. The data included only the latest 1000 posts from each SMI. The data includes profile information, photos, captions, number of likes, tags associated with photos, and users’ social network that includes friends and followers.

Statistics

The statistics of various data features were extracted using simple maths such as averages, percentages and totals. This data is used for the correlation analysis in the next stages.

NLP Models

The data had to be pre-processed before text-mining techniques could be applied to extract the language aspects of the posts. For this purpose, sentence tokenization, word tokenization and normalization techniques were employed (Manning, Raghavan, & Schütze, 2008). A text is structured on several levels – a document consists of paragraphs which is composed of sentences, which are sequences of words. Lemmatization is the process of reducing the dimensions of the data for better performance. The lemma rule generates a base from a word by stripping some prefix and suffix and prepending and appending new prefix and suffix. For example, words, cats and happiest are reduce to cat and happy respectively. Text pre-processing involves the removal of short sentences, punctuations, tokenization and lemmatization. This process also filtered out posts that are fewer than three sentences in length so that at least three segments of posts are available for the narrative analysis.

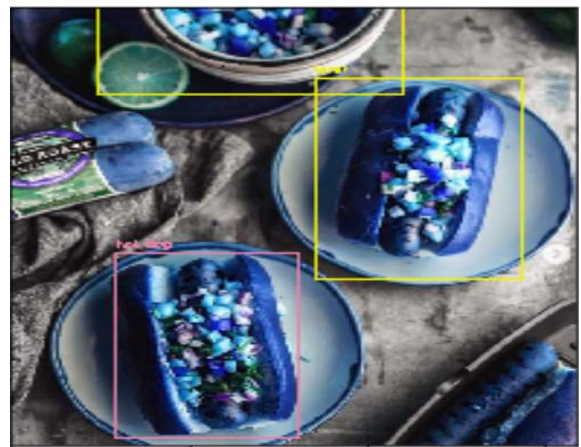
Image Extraction Models

Instagram API extracts the URLs of the images for each post. We observe that some posts have more than one image and our task is to extract the objects in the images. Image extraction involves two sub-tasks; Image detection and Image Identification. In our solution implementation we used OpenCV and cvlib, open source computer vision library that is used in real-time computer vision (Redmon et al., 2016). For this study, an image-detection python algorithm based on cvlib – a high-level library for object detection – was built. cvlib uses a model pre-trained on the COCO dataset, built on YOLO model and is capable of detecting 80 common objects. These objects include people, vehicles, animals, kitchen items, food etc. YOLO passes the n by n image only once in a fully convolutional neural network (FCNN), which makes it quite fast and real time. It splits the image into grids of dimension m by m and generates bounding boxes and their class probabilities. An example of image extraction is shown in the Figure 2. The rectangular boxes indicate the image detection part (Figure 2a) and in this example, we observe rectangles indicating three objects detected. Image identification indicates labels on the top of the rectangular boxes (Figure 2b). In Figure 2b, we observe three labels; bowl, hotdog, hotdog. Moreover, we noticed that lemon is not detected. When this process is applied to Instagram posts, it enables the identification of the protagonists in the Instagram pictures.

FIGURE 2
IMAGE DETECTION AND IDENTIFICATION



(a) Jpeg image input to YOLO image classifier



(b) Rectangular boxes to indicate detection and labels indicate identification

Text Mining and LIWC Tools

In this stage, we have multiple tasks for mining the insights from the data. Top words refer to the most common words that appear in the posts after filtering stopwords such as prepositions, conjunctions etc. The stopwords are removed using a lexicon of the most frequent words in English language (Manning et al., 2008). We applied basic text-mining techniques and word token distribution using python programming to extract the top words and hashtags from the posts. Subsequently, we used LIWC APIs to extract the emotions, topics and language aspects. The outputs from this stage enabled the generation of emotional arcs as well as the analysis of protagonists in the Instagram posts. Based on the emotions extracted for each text segment of the post, the graphical arc is created for visualizations. We converted all the post from the previous stage into a story arc. These arcs are then automatically classified into the five categories in the next task.

To identify the story's protagonist, the posts were analysed for the prevalence of first, second, and third-person pronouns. Third-person pronouns are associated with an observer's perspective while second-person pronouns support an actor's perspective. First-person pronoun effects depend on the situational context (Brunye et al., 2009; Brunye et al., 2011).

Spearman's Rank Correlations

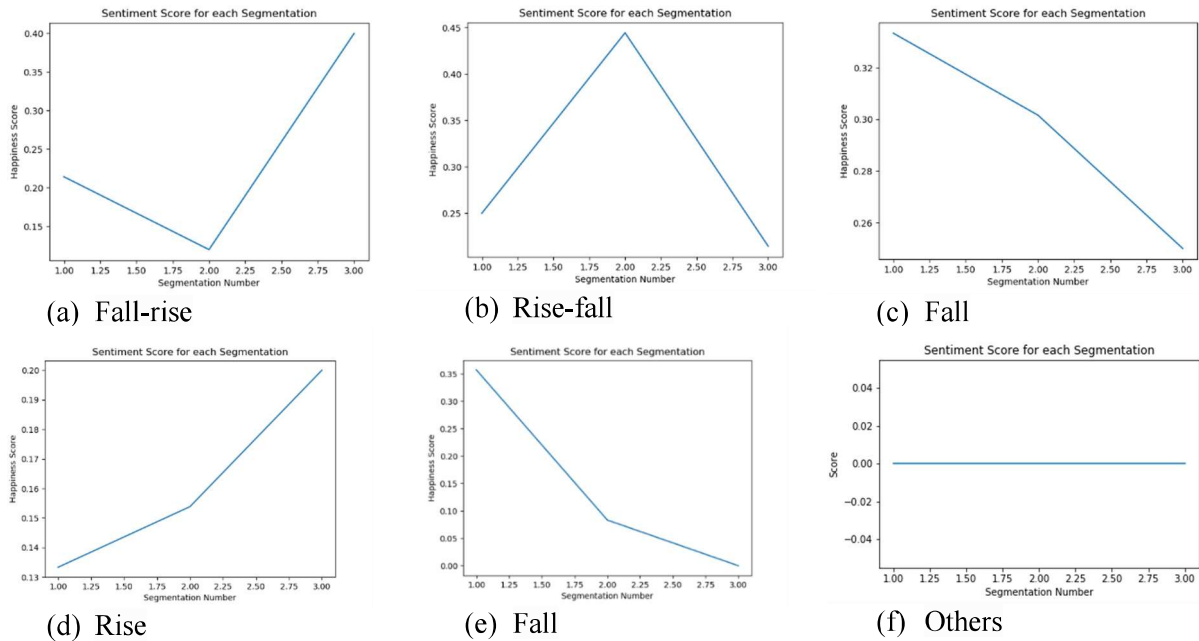
The Spearman correlation evaluates the monotonic relationship between two continuous or ordinal variables. The Spearman correlation coefficient is based on the ranked values for each variable rather than the raw data. Spearman's rank correlation (Spearman, 1904), equals the Pearson correlation computed from the ranks of the observations. In our case, analyze audience engagement with SMIs' posts is measured by the number of likes. This metric is suitable for measuring relationships with ordinal variables such as 'number of likes'. LIWC produces scores for each category of language and emotion aspects which is a continuous variable. Spearman correlation is used to evaluate relationship between number of likes and the scores of the language aspects. For example, we can it to evaluate whether the audience engagement (order of number of likes) is related to the total number of emotions in the post.

Image-Matching Models

To label the arcs, we tried two different techniques; image classification and template matching. The fundamental task of image classification is to categorise the images into specific classes or groups. The various applications such as vehicle navigation, robot navigation and remote sensing by using image classification technology. In our case, we have five classes of story arcs. Keras framework is a deep learning based convolution network model used to build image classifiers (Raschka & Mirjalili, 2017). We manually labelled 200 arcs into various categories; fall, fall-rise, rise, rise-fall, and others. In our preliminary experiments this model fails as the graphs are lines and CNN fails to detect the full shape in the image. The accuracy was only 50%. Therefore, we used the model, template matching, which can be useful to match the geometrical shapes or the given templates.

Template matching is a method for searching and finding the location of a template image in a larger image (Mordvintsev & Abid, 2013). OpenCV comes with a function – i.e. `cv2.matchTemplate` – for this purpose. It simply slides the template image over the input image (as in 2D convolution) and compares the template and patch of input image. We tried various image comparison methods between the trained images and remaining test data story arc images; `TM_CCORR_NORMED`, `TM_COEFF_NORMED`, `TM_SQDIFF_NORMED`. We choose the labelled arc with the lowest matching score to label the arcs in the test data set. We took a small dataset and `TM_SQDIFF_NORMED` method provided 95% accuracy. Figure 3 shows examples of story arcs from the posts.

FIGURE 3
STORY ARCS CREATED USING EMOTION SCORES FROM THE LIWC API ON
SEGMENTS OF POSTS



Subsequently, the image matching model based on template matching method was used to auto-label the remaining 8600 arcs. This enabled us to analyse the influencer’s story arcs in Instagram posts.

FINDINGS AND ANALYSIS

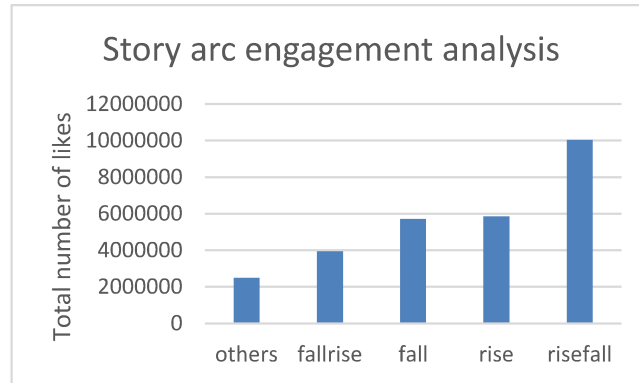
The most common story arc in SMI’s Instagram posts is the ‘rise-fall’ arc, followed by the ‘fall’ and the ‘rise’ arcs (see Table 2). The below table answers our RQ1. The rise-fall and fall are the most prevalent story arcs used by the influencers in Instagram posts.

TABLE 2
PREVALENCE OF TYPES OF STORY ARCS

Arc Type	% of Posts
Rise-fall	35.96
Fall	22.53
Rise	19.03
Fall-rise	14.21
Others	8.28

In addition, the story arc engagement statistics show that the ‘rise-fall’ arc, followed by the ‘rise’ and ‘fall’ arcs, elicit the highest number of likes (see Figure 4) which answers our RQ2, emotional arcs which trigger audience engagement.

**FIGURE 4
STORY ARC ENGAGEMENT STATISTICS**



To investigate whether SMIs positioned themselves, their audiences or the brands they endorse as the protagonist of their stories, we analyzed the use of hashtags, personal pronouns and images in their Instagram posts. Table 3 shows the personal pronoun statistics in the posts.

**TABLE 3
PROTAGONISTS IN THE POSTS**

Pronoun	Appearance in posts
First-person plural	38.99%
First-person singular	7.41%
Second-person	14.47%
Third-person plural	1.88%
Third-person singular	2.57%

It can be seen that first-person plural pronouns (see Figure 3a) are the dominant form of pronoun, suggesting that SMIs mostly positioned themselves as the protagonists in their brand stories. This is supported by image analysis, which shows that “person” is the most common object in the Instagram images (see table 4). “Person” is followed by objects that are relevant to the topics of interest for each SMI, such as “cake”, “sandwich” and “bowl” for the food influencer “Littlemissbento”. In addition, the hashtags are dominated by the names of brands (see Figure 3b). Therefore, the first hypothesis – i.e. SMIs more frequently position their audiences (rather than themselves or the brands they endorse) as the protagonists in their stories – is not supported.

**TABLE 4
TOP FIVE OBJECTS IN SELECTED SMI’S INSTAGRAM IMAGES**

User	Top words in posts	Top objects in images
Nicolechangmin	skin', 'product', 'look', 'free', 'back'	person , cup , dog , bowl , cat
Melissackoh	favourite', 'home', 'best', 'beautiful', 'll'	person , potted , plant , cake , vase
Wafflecreamies	look', 'back', 'work', 'take', 'travel'	person , chair , tie , bottle , car
Littlemissbento	bento', 'littlemissbento', 'cute', 'bread', 'make'	cake , person , sandwich , bowl, table
naomineo_	guy', 'feel', 'giveaway', 'anyway', 'back'	person , cell , phone , chair , umbrella

Cheekiemonkies	kid', 'link', 'bio', 'monkie', 'cheekiemonkie'	person , toilet , cell , phone , chair
Cottonthemaltese	cotton', 'dog', 'hooman', 'friend', 'treat'	dog , book , tie , person , cake
Amebeverly	back', 'destination', 'world', 'link', 'bio'	person , cup , chair , handbag , potted

To further understand audience engagement, we analyzed the correlations between various aspects of the Instagram posts – i.e. personal pronouns, emotions, and social words– and the number of ‘likes’ (see Table 6). The posts of the top five SMIs show positive correlations between the prevalence of personal pronouns, positive emotions, and social words and the number of ‘likes’; the correlations are either positive (20% - 50%) or strongly positive (50% and above). The converse is true for the bottom five SMIs. Thus, the second and third hypotheses are supported.

TABLE 5
CORRELATIONS BETWEEN THE NUMBER OF ‘LIKES’ AND LANGUAGE ASPECTS FOR THE TOP FIVE AND BOTTOM FIVE INFLUENCERS

Names	Rank	Category	I	You	Emotions	PosEmo	NegEmo	Social
Cheekiemonkies	Top 5	Parenting	0.68	0.58	0.43	0.55	0.06	0.61
Amebeverly	Top 5	Travel	0.64	0.24	0.04	0.01	0.32	0.68
cottonthemaltese	Top 5	Pets	0.58	0.33	0.24	0.31	0.11	0.48
Nicolechangmin	Top 5	Beauty	0.45	0.29	0.13	0.18	0.24	0.44
Melissackoh	Top 5	Fashion	0.42	0.38	0.44	0.49	0.00	0.69
hannah_chia	Bottom 5	Food	-0.04	-0.28	-0.27	-0.31	0.01	-0.16
Ieatishootipost	Bottom 5	Food	-0.09	0.11	0.20	0.20	-0.07	-0.02
Seowie	Bottom 5	Fitness	-0.20	0.35	-0.31	-0.35	-0.05	-0.50
Dreachong	Bottom 5	Fashion	-0.36	-0.06	-0.15	-0.23	0.30	-0.17
Sophiewillocq	Bottom 5	Travel	-0.41	-0.01	0.07	-0.02	0.17	-0.09

DISCUSSION

Theoretical Implications

While the role of narrative transportation in advertising has been extensively studied (e.g. Escalas, 2007; Green, 2006), there have been extremely few studies to date on its role in influencer marketing. This study is one of the very first to use data mining techniques to understand SMI narrative persuasion on Instagram. Like Feng, Chen and Kong (2020), we posit that SMIs’ Instagram posts are exemplars of narrative persuasion and show that narrative persuasion theory can aid in the understanding of how narratives may be effective in the context of influencer marketing.

This study also confirms previous findings on the link between the use of first-person perspectives as well as positive emotions on behavioural outcomes (e.g. higher engagement), and extends these findings to social networking sites such as Instagram. The first-person storytelling that is found in SMIs’ Instagram posts helps foster follower identification with the SMIs, which may lead to narrative persuasion (Hoeken et al., 2016). Moreover, since the majority of SMI posts in this study are associated with positive emotions, the vicarious transfer of these emotions through character identification is likely to generate positive affect in the followers (Escalas, 2007; Green, 2006). The positive affect experienced by the followers may in turn promote behaviors such as “likes” on social media (Hong, Chen & Li, 2017).

Practical Implications

This study has a number of practical implications. First, it shows that storytelling – especially first-person storytelling – by SMIs exerts an influence on audience engagement. Hence, SMIs and the brands that collaborate with them could adopt a more concerted narrative approach to influencer marketing. While the relationship between SMI-category fit and follower engagement is not a focus of this study, the literature shows that tight alignment between the SMI's domain expertise and the category (e.g. food, travel, parenting) that he/she belongs to is crucial to the success of influencer marketing (Backaler, 2018).

Second, this study challenges conventional storytelling advice (e.g. Adi et al., 2015; Sachs, 2012), which posits that organizations and brands should place the consumer in the role of the story protagonist. In this study, the consumers of the SMI as protagonist and its effectiveness in eliciting audience response both suggest that consumers not only welcome– but may in fact expect – to learn about brands (and other topics) from the SMI's perspective. The desire to foster parasocial relationships with online celebrities may be precisely why they became SMI followers in the first place (Chung & Cho, 2017). In addition, when an SMI features himself/herself as the story protagonist and accompanies the story with a photo of himself/herself (which happened in the majority of the Instagram posts), he/she adds an authentic voice to his/her first-person brand narrative (Feng et al., 2020). Hence, SMIs would do well to tell stories using the first-person perspective and reinforce it with accompanying photos of themselves.

Finally, SMIs should infuse their posts with more positive emotions, do it more frequently, and avoid negative emotions if their aim is to engage their followers. Previous studies (e.g. Berger & Milkman, 2012; Vosoughi, Roy & Aral, 2018) have shown a positive relationship between high-arousal emotion in content and content virality. Moreover, high-arousal content (in both positive and negative form) is more memorable than low-arousal content – this is in line with research on norm violations (Nelson-Field, 2020). While norm-violating content (e.g. content that is considered offensive) may enhance virality and memorability, its risk to brand reputation needs to be seriously considered, especially as brand messages typically feature positive emotional appeals (Nelson-Field, 2020).

Limitations and Suggestions for Future Research

The dominance of the 'rise-fall' arc and the fact that it triggered the highest number of 'likes' is a surprising finding, as stories tend to end on an emotionally high (versus an emotionally low or falling) note (Hambry & Brinberg, 2016). Nonetheless, the fact that the largest percentage of SMI posts ends on a falling note may have something to do with our removal of text that comprised only hashtags. When hashtags in the final segment of a post are removed, and these hashtags have positive emotions, their removal could inadvertently result in a falling arc.

There is a key limitation in our image analysis and topic extraction. In terms of image analysis, we were able to extract only the objects but not the actions of the protagonist in the image. Action analysis is important, as portrayed action has been found "to affect viewer likelihood to comment on the visualized narrative" (Farace, van Laer, Ruyter, & Wetzels, 2017, p. 1975). In addition, we were unable to extract more specific topics such as "travel", "parenting" and "pets". Instead, we grouped "travel" and "pets" under "leisure". Similarly, the topic "parenting" is grouped under the "social" category. While words such as "baby" and "child" can be found in the "social" category, some words that are relevant to parenting (e.g. "care", "needs", "sleep") are missing – the absence of these words can impact the accuracy of findings.

Future studies can take the form of sentiment analysis of SMIs' Instagram images and hashtags and their impact on audience engagement with brand stories.

CONCLUSION

The rising importance of SMIs in brand communication is happening at the same time that storytelling is gaining prominence as a promotional tactic. To date, few studies have focused on SMIs' brand storytelling on social media by the influencers. This study is a preliminary effort to fill that literature gap through the text-mining and analysis of SMIs' Instagram posts.

REFERENCES

- Abidin, C. (2015). Communicative < 3 intimacies: Influencers and perceived interconnectedness. *Ada*, 8. Retrieved from <https://adanewmedia.org/2015/11/issue8-abidin/>
- Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86-100.
- Adaval, R., & Wyer, R.S. (1998). The role of narratives in consumer information processing. *Journal of Consumer Psychology*, 7(3), 207-45.
- Adi, A., Crisan, C., & Dinca, R.C. (2015). Stories, heroes and commercials. Spreading the message across with a new type of responsibility. *Management Dynamics in the Knowledge Economy*, 3(4), 749-764.
- Argo, J.J., Zhu, R., & Dahl, D.W. (2008). Fact or fiction: An investigation of empathy differences in response to emotional melodramatic entertainment. *Journal of Consumer Research*, 34(5), 614-23.
- Backaler, J. (2018). *Digital influence. Unleash the power of influencer marketing to accelerate your global business*. Cham, Switzerland: Palgrave Macmillan.
- Bakker, D. (2018). Conceptualizing influencer marketing. *Journal of Emerging Trends in Marketing and Management*, 1(1), 79–87.
- Banerjee, S.C., & Green, M.C. (2012). “I quit” versus “I’m sorry I used”: A preliminary investigation of variations in narrative ending and transportation. *Psychology and Health*, 27(11), 1308-22.
- Belton, A. (2019). *Social influence: Leveraging the power of influencers to build brand advocacy*. Association of National Advertisers. Retrieved from <https://www.ana.net/blogs/show/id/mm-blog-2018-12-leveraging-the-power-of-influencers-to-build-brand-advocacy>
- Bergen, B., & Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. In J.O. Ostman & M. Fried (Eds.), *Construction grammars: Cognitive grounding and theoretical extensions* (pp. 147-190). Amsterdam, Netherlands: John Benjamins Publishing Company.
- Berger, J., & Milkman, K.L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192-205.
- Berger, J., & Schwartz, E. (2011). What drives immediate and ongoing word of mouth? *Journal of Marketing Research*, 48(5), 869-80.
- Brunye, T.T., Ditman, T., Mahoney, C.R., & Taylor, H.A. (2011). Better you than I: Perspectives and emotion simulation during narrative comprehension. *Journal of Cognitive Psychology*, 23(5), 659-666.
- Brunye, T.T., Ditman, T., Mahoney, C.R., Augustyn, J.S., & Taylor, H.A. (2009). When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. *Psychological Science*, 20(1), 27-32.
- Busselle, R., & Bilandzic, H. (2008) Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, 18 (2), 255–80.
- Carter, D. (2016). Hustle and brand: The sociotechnical shaping of influence. *Social Media + Society*, 2(3), 1-12.
- Chatzigeorgiou, C. (2017). Modelling the impact of social media influencers on behavioral intentions of millennials: The case of tourism in rural areas in Greece. *Journalism of Tourism, Heritage & Services Marketing*, 3(2), 25-29.
- Chung, S., & Cho, H. (2017). Fostering parasocial relationships with celebrities on social media: Implications for celebrity endorsement. *Psychology & Marketing*, 34(4), 481-495.
- De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, 36(5), 798–828.
- De Vries, L., Gensler, S., & LeeFlang, P.S.H. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. *Journal of Interactive Marketing*, 26(2), 83-91.

- Dittman, T., Brunye, T.T., Mahoney, C.R., & Taylor, H. (2010). Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. *Cognition*, *115*(1), 172-78.
- Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. *Computers in Human Behavior*, *68*, 1-7.
- Eagly, A.H., & Chaiken, S. (1993). *The psychology of attitudes*. Boston, Massachusetts: Academic Press.
- Erz, A., & Christensen, A-B.H. (2018). Transforming consumers into brands: Tracing transformation processes of the practice of blogging. *Journal of Interactive Marketing*, *43*, 69-82.
- Escalas, J.E. (2004a). Imaging yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising*, *33*(2), 37-48.
- Escalas, J.E. (2004b). Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology*, *14*(1/2), 168-80.
- Escalas, J.E. (2007). Self-referencing and persuasion: Narrative transportation versus analytical elaboration. *Journal of Consumer Research*, *33*(4), 421-29.
- Escalas, J.E., Moore, M.C., & Britton, J.E. (2004). Fishing for feelings? Hooking viewers helps! *Journal of Consumer Psychology*, *14*(1-2), 105-14.
- Farace, S., van Laer, T., de Router, K., & Wetzels, M. (2017). Assessing the effect of narrative transportation, portrayed action, and photographic style on the likelihood to comment on posted selfies. *European Journal of Marketing*, *51*(11/12), 1961-1979.
- Feehan, B. (2018, April 2). *2018 social media industry benchmark report*. RivalIQ. Retrieved from <https://www.rivaliq.com/blog/2018-social-media-industry-benchmark-report/>
- Feng, Y., Chen, H., & Kong, Q. (2020). An expert with whom I can identify: the role of narratives in influencer marketing. *International Journal of Advertising*. <https://doi.org/10.1080/02650487.2020.1824751>
- Freberg, K., Graham, K., McGaughey, K., & Freberg, L. (2011). Who are the social media influencers? A study of public perceptions of personality. *Public Relations Review*, *37*(1), 90-92.
- Gerrig, R.J. (1999). *Experiencing narrative worlds*. Abingdon, UK: Routledge.
- Green, M.C. (2006). Narratives and cancer communication. *Journal of Communication*, *56*(S1), S163-83.
- Green, M.C. (2008). Research challenges: Research challenges in narrative persuasion. *Information Design Journal*, *16*(1), 47-52.
- Green, M.C., & Brock, T.C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701-21.
- Green, M.C., & Brock, T.C. (2002). In the mind's eye: Transportation imagery model of narrative persuasion. In M.C. Green, J.J., Strange, & T.C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 315-341). New York, NY: Lawrence Erlbaum Associates Publishers.
- Green, M.C., & Donahue, J.K. (2011). Persistence of belief change in the face of deception: The effect of factual stories revealed to be false. *Media Psychology*, *14*(3), 312-31.
- Hambry, A., & Brinberg, D. (2016). Happily ever after: How ending valence influences narrative persuasion in cautionary stories. *Journal of Advertising*, *45*(4), 498-508.
- Hartung, F., Burke, M., Hagoort, P., & Willems, R.M. (2015). Taking perspective: Personal pronouns affect experiential aspects of literary reading. *PLoS ONE*, *11*(5), e0154732.
- Heath, C., & Heath, D. (2007). *Made to stick. Why some ideas survive and others die*. New York, NY: Random House.
- Hoeken, H., Kolthoff, M., & Sanders, J. (2016). Story perspective and character similarity as drivers of identification and narrative persuasion. *Human Communication Research*, *42*(2), 292-311.
- Hoeken, H., Kolthoff, M., & Sanders, J. (2016). Story perspective and character similarity as drivers of identification and narrative persuasion. *Human Communication Research*, *42*(2), 292-311.
- Hollebeck, L.D. (2011). Exploring customer brand engagement: Definition and themes. *Journal of Strategic Marketing*, *19*(7), 555-573.
- Hong, C., Chen, Z.F., & Li, C. (2017). "Liking" and being "liked": how are personality traits and demographics associated with giving and receiving "likes" on Facebook? *Computers in Human Behavior*, *68*, 292-299.

- Hughes, C., Swaminathan, V., & Brooks, G. (2019). Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns. *Journal of Marketing*, 83(5), 78-96.
- Kahneman, D. (2013). *Thinking, fast and slow*. New York, NY: Farrar, Straus & Giroux.
- Kim, E., Ratneshwar, R., & Thorson, E. (2017). Why narrative ads work: An integrated process explanation. *Journal of Advertising*, 46(2), 283-96.
- Kim, M., Lee, J.K., & Lee, K-Y. (2019). Interplay of content type and product type in the consumer response to native advertising on social media. *Asian Journal of Communication*, 29(6), 464-482.
- Lim, H. (2019). *Top Instagram influencers in Singapore you may consider hiring*. sbo. Retrieved from <https://sbo.sg/guide/top-instagram-influencers-singapore/>
- Lin, H.C., Bruning, P.F., & Swarna, H. (2018). Using online opinion leaders to promote the hedonic and utilitarian value of products and services. *Business Horizons*, 61(3), 431-442.
- Manning, C., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. New York, NY: Cambridge University Press.
- Marwick, A.E. (2015). Instafame: Luxury selfies in the attention economy. *Public Culture*, 27, 137-160.
- Mediakix. (2018, March 6). *The influencer industry global ad spend: A \$5 – 10 billion market by 2020*. Retrieved from <https://mediakix.com/blog/influencer-marketing-industry-ad-spend-chart/>
- Mordvintsev, A., & Abid, K. (2013). *Opencv-python tutorials*. Retrieved from <http://opencv-python-tutroals.readthedocs.io>
- Nelson-Field, K. (2020). *The attention economy and how media works*. Palgrave Macmillan.
- Oatley, K. (1999). Meeting of minds: Dialogue, sympathy and identification in reading fiction. *Poetics*, 26(5-6), 439-54.
- Papeo, L., Corradi-dell'acqua, C., & Rumiati, R.I. (2011). “She” is not like “I”: The tie between language and action is in our imagination. *Journal of Cognitive Neuroscience*, 23(12), 3939-48.
- Pennebaker, J.W., & Graybeal, A. (2001). Patterns of natural language use. Disclosure, personality, and social integration. *Current Directions in Psychological Science*, 10(3), 90–93.
- Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., & Booth, R. (2007). *The development and psychometric properties of LIWC2007*. Austin, Texas: LIWC.net.
- Petty, R.E., & Cacioppo, J.T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Berlin, Germany: Springer.
- Phillips, B.J., & McQuarrie, E.F. (2010). Narrative and persuasion in fashion advertising. *Journal of Consumer Research*, 37(3), 368-92.
- Raschka, S., & Mirjalili, V. (2017). *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow*. Birmingham, UK: Packt Publishing.
- Reagan, A.J., Mitchell, L., Kiley, D., Danforth, C.M., Dodds, P.S. (2016). The emotional arcs of stories are dominated by six shapes. *EPJ Data Science*, 5(31), 1-12.
- Redmon, J., Divvala, S.K., Girshick, R.B., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788.
- Sachs, J. (2012). *Winning the story wars: Why those who tell (and live) the best stories will rule the future*. Brighton, Massachusetts: Harvard Business Review Press.
- Sanford, A.J., & Emmott, C. (2012). *Mind, brain and narrative*. Cambridge, UK: Cambridge University Press.
- Scott, K. (2015). The pragmatics of hashtags: Inference and conversational style on Twitter. *Journal of Pragmatics*, 81, 8-20.
- Scott, K. (2018). ‘Hashtags work everywhere’: The pragmatic functions of spoken hashtags. *Discourse, Context & Media*, 22, 57-64.
- Senft, T.M. (2013). Microcelebrity and the branded self. In J. Hartley, J. Burgess, & A. Bruns (Eds.), *A companion to new media dynamics* (pp. 346-354). Chichester, UK: Blackwell Publishing.
- Slater, M.D., & Rouner, D. (2012). Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory*, 12(2), 173-91.

- Song, S.Y., Cho, E., & Kim, Y-K. (2017). Personality factors and flow affecting opinion leadership in social media. *Personality and Individual Differences, 114*, 16-23.
- Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15*, 201–293.
- Stern, B.B. (1994). Authenticity and the textual persona: Postmodern paradoxes in advertising narrative. *International Journal of Research in Marketing, 11*(4), 387-400.
- Sudha, M., & Sheena, K. (2017). Impact of influencers in consumer decision process: The fashion industry. *SCMS Journal of Indian Management, 14*(3), 14-30.
- Sundermann, G., & Rabe, T. (2019). Strategic communication through social media influencers. Current state of research and desiderata. *International Journal of Strategic Communication, 13*(4), 278-300.
- 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.
- van Laer, T., De Ruyter, K., Visconti, L.M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research, 40*(5), 797-817.
- van Laer, T., Escalas, J.E., Ludwig, S., & van den Hende, E.A. (2019). What happens in Vegas stays on TripAdvisor? A theory and technique to understand narrativity in consumer reviews. *Journal of Consumer Research, 46*(2), 267-285.
- Voorveld, H.A.M. (2019). Brand communication in social media: A research agenda. *Journal of Advertising, 48*(1), 14-26.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151.
- Wang, J., & Calder, B.J. (2006). Media transportation and advertising. *Journal of Consumer Research, 33*(2), 151–62.
- Wang, J., & Calder, B.J. (2009). Media engagement and advertising: Transportation, matching, transference, and intrusion. *Journal of Consumer Psychology, 19*(3), 546–55.
- Wu, T. (2017). *The attention merchants. The epic scramble to get inside our heads*. New York, NY: Vintage.
- Zak, P. (2015). Why inspiring stories make us react: The neuroscience of narrative. *Cerebrum, 2*. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4445577/>