

Syracuse University

**SURFACE**

---

Dissertations - ALL

SURFACE

---

May 2018

## Microcelebrity Practices: A Cross-Platform Study Through a Richness Framework

Sikana Tanupabrungsun  
*Syracuse University*

Follow this and additional works at: <https://surface.syr.edu/etd>



Part of the [Social and Behavioral Sciences Commons](#)

---

### Recommended Citation

Tanupabrungsun, Sikana, "Microcelebrity Practices: A Cross-Platform Study Through a Richness Framework" (2018). *Dissertations - ALL*. 879.

<https://surface.syr.edu/etd/879>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

## ABSTRACT

Social media have introduced a contemporary shift from broadcast to participatory media, which has brought about major changes to the celebrity management model. It is now common for celebrities to bypass traditional mass media and take control over their promotional discourse through the practice of *microcelebrity*.

The theory of microcelebrity explains how people turn their public persona into media content with the goal of gaining and maintaining audiences who are regarded as an aggregated fan base. To accomplish this, the theory suggests that people employ a set of online self-presentation techniques that typically consist of three core practices: identity constructions, fan interactions and promoting visibility beyond the existing fan base. Studies on single platforms (e.g., Twitter), however, show that not all celebrities necessarily engage in all core practices to the same degree. Importantly, celebrities are increasingly using multiple social media platforms simultaneously to expand their audience, while overcoming the limitations of a particular platform. This points to a gap in the literature and calls for a cross-platform study.

This dissertation employed a mixed-methods research design to reveal how social media platforms i.e., Twitter and Instagram, helped celebrities grow and maintain their audience. The first phase of the study relied on a richness scoring framework that quantified social media activities using *affordance richness*, a measure of the ability of a post to deliver the information necessary in affording a celebrity to perform an action by using social media artifacts. The analyses addressed several research questions regarding social media uses by different groups of celebrities and how the audience responded to different microcelebrity strategies. The findings informed the design of the follow-up interviews with audience members. Understanding

expectations and behaviors of fans is relevant not only as a means to enhance the practice's outcome and sustain promotional activity, but also as a contribution to our understandings about contemporary celebrity-fans relationships mediated by social media.

Three findings are highlighted. First, I found that celebrities used the two platforms differently, and that different groups of celebrities emphasized different core practices. This finding was well explained by the interviews suggesting that the audiences had different expectations from different groups of celebrities. Second, microcelebrity strategies played an important role in an audience's engagement decisions. The finding was supported by the interviews indicating that audience preferences were based on some core practices. Lastly, while their strategies had no effect on follow and unfollow decisions, the consistency of the practices had significant effects on the decisions.

This study makes contributions to the theory of Microcelebrity and offers practical contributions by providing broad insights from both practitioners' and audiences' perspectives. This is essential given that microcelebrity is a learned practice rather than an inborn trait.

MICROCELEBRITY PRACTICES: A CROSS-PLATFORM STUDY THROUGH A  
RICHNESS FRAMEWORK

by

Sikana Tanupabrungsun

B.Eng., King Mongkut's University of Technology Thonburi, Thailand, 2008

M.Eng., King Mongkut's University of Technology Thonburi, Thailand, 2012

Dissertation

Submitted in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Information Science and Technology.

Syracuse University

May 2018

Copyright © Sikana Tanupabrungsun 2018  
All Rights Reserved

## **ACKNOWLEDGEMENT**

I would like to express my sincere gratitude to my doctoral advisor, Jeff Hemsley, for his endless encouragement, mentorship and support of my study, research and my life in general. I could not imagine having a better doctoral advisor. I am also grateful to the members of my committee for their insightful comments and thoughtful feedback throughout writing this dissertation. This dissertation would not have been possible without funding from the Center for Computational and Data Sciences (CCDS).

Last but not least, I would like to thank my family and friends for their encouragement and support during the hard times.

# TABLE OF CONTENTS

ABSTRACT.....	I
ACKNOWLEDGEMENT .....	V
TABLE OF CONTENTS.....	VI
LIST OF FIGURES .....	VIII
LIST OF TABLES .....	X
LIST OF EQUATIONS .....	XIII
CHAPTER 1 INTRODUCTION .....	1
1.1 BACKGROUND AND MOTIVATION .....	2
1.2 RESEARCH PROBLEM .....	4
1.3 RESEARCH METHODS AND FINDINGS .....	5
1.4 TERMINOLOGIES .....	9
1.5 CONTRIBUTIONS .....	10
1.6 DOCUMENT ORGANIZATION .....	12
CHAPTER 2 LITERATURE REVIEW .....	13
2.1 CELEBRITY STUDIES .....	13
2.1.1 <i>Theory of Microcelebrity</i> .....	14
2.1.2 <i>Core Practices of Microcelebrity</i> .....	18
2.1.3 <i>Related Studies</i> .....	20
2.1.4 <i>Summary</i> .....	26
2.2 THEORETICAL MODELS .....	28
2.2.1 <i>Theory of Affordances</i> .....	29
2.2.2 <i>Media Richness Theory</i> .....	38
2.2.3 <i>Summary</i> .....	52
2.3 TECHNOLOGY ARTIFACTS, ENGAGEMENT AND FOLLOWERSHIP ON SOCIAL MEDIA.....	54
2.3.1 <i>Technological Artifacts</i> .....	55
2.3.2 <i>Followership on Social Media</i> .....	57
2.3.3 <i>Follower Engagement on Social Media</i> .....	64
2.3.4 <i>Summary</i> .....	69
2.4 LITERATURE SUMMARY .....	69
2.5 RESEARCH QUESTIONS .....	70
CHAPTER 3 METHODOLOGY OVERVIEW .....	74
3.1 METHODOLOGY REVIEW .....	75
3.2 QUANTITATIVE STUDIES.....	76
3.3 QUALITATIVE STUDY.....	77
CHAPTER 4 QUANTITATIVE METHODS .....	78
4.1 RICHNESS FRAMEWORK.....	78
4.1.1. <i>Structural Design</i> .....	80
4.1.2. <i>Scoring Development</i> .....	83
4.2. SCORE ANALYSIS .....	121

4.2.1 Dataset .....	121
4.2.2 Comparisons by platforms .....	122
4.2.3 Comparison by celebrity types .....	123
4.2.4 Audience response .....	123
4.3 SUMMARY .....	125
CHAPTER 5 QUANTITATIVE RESULTS AND FINDINGS .....	127
5.1 COMPARISONS BY PLATFORMS .....	127
5.2 COMPARISON BY CELEBRITY TYPES .....	129
5.3 AUDIENCE RESPONSE .....	131
<i>Model 1: Engagement scores and richness measures</i> .....	133
<i>Model 2: Mean engagement scores and variance of richness scores</i> .....	147
<i>Model 3: Follower changes and mean richness scores</i> .....	162
<i>Model 4: Follower changes and variances in richness scores</i> .....	170
5.4. SUMMARY .....	178
CHAPTER 6 QUALITATIVE METHODS .....	182
6.1 STUDY DESIGN .....	182
6.1.1 Participant Recruitment .....	183
6.1.2 Interviews .....	186
6.2 QUALITATIVE CODING .....	189
6.3 SUMMARY .....	198
CHAPTER 7 QUALITATIVE RESULTS AND FINDINGS .....	199
7.1 FINDINGS .....	199
7.1.1 Microcelebrity Practices on Different Platforms and Different Celebrity Types .....	201
7.1.2 The Effects of Microcelebrity Strategies on Maintaining and Growing an Audience .....	203
7.1.3 The Effects of the Consistency in Microcelebrity Strategies on Maintaining and Growing an Audience .....	205
7.2 SUMMARY .....	206
CHAPTER 8 .....	209
DISCUSSION .....	209
8.1 RESEARCH METHODS .....	210
8.2 THE FRAMEWORK AND NOTION OF AFFORDANCE RICHNESS .....	212
8.3 FINDINGS .....	215
8.3.1 Identity Construction .....	216
8.3.2 Fans Interaction .....	219
8.3.3 Visibility Promotion .....	221
8.4 MICROCELEBRITY AS A LEARNED TECHNIQUE .....	223
8.5 LIMITATIONS AND FUTURE WORK .....	225
CHAPTER 9 .....	228
CONCLUSION .....	228
APPENDICES .....	232
RECRUITING LETTER .....	232
INTERVIEW PROTOCOL .....	233
CODEBOOK .....	237
IRB EXEMPTION .....	242
REFERENCES .....	243
VITA .....	269



## LIST OF FIGURES

<b>Figure 1.1</b> Overall Research Scheme.....	8
<b>Figure 3.1</b> Research Paradigm (adapted from (Creswell, 2013)).....	76
<b>Figure 4.1</b> Distributions of tweet annotations (Rich vs. Lean) by AMT workers.....	92
<b>Figure 4.2</b> Distributions of Instagram post annotations by AMT workers.....	94
<b>Figure 4.3</b> Model training process.....	97
<b>Figure 5.1</b> Normal plots of the differences between pairs in each richness dimension.....	128
<b>Figure 5.2</b> Normal plots of the samples in each richness dimension for each platform.....	130
<b>Figure 5.3</b> Normal plot of the residuals .....	136
<b>Figure 5.4</b> Normal plot of the residuals .....	139
<b>Figure 5.5</b> Normal plot of the residuals .....	142
<b>Figure 5.6</b> Normal plot of the residuals .....	145
<b>Figure 5.7</b> Normal plot of the residuals .....	151
<b>Figure 5.8</b> Normal plot of the residuals .....	154
<b>Figure 5.9</b> Normal plot of the residuals .....	157
<b>Figure 5.10</b> Normal plot of the residuals .....	160
<b>Figure 5.11</b> Normal plot of the residuals .....	165

<b>Figure 5.12</b> Normal plot of the residuals .....	168
<b>Figure 5.13</b> Normal plot of the residuals .....	173
<b>Figure 5.14</b> Normal plot of the residuals .....	176
<b>Figure 6.1</b> The structure of the codebook for analyzing interviews data.....	197

## LIST OF TABLES

<b>Table 2.1</b> Media richness and traditional media.....	41
<b>Table 2.2</b> Kaplan and Haenlein’s (2010, 2011) classification scheme.....	50
<b>Table 4.1</b> Statistics of frequency per user, grouped by celebrity types and platforms.....	88
<b>Table 4.2</b> The numbers of instances (%) labelled as <i>rich</i> in each dimension.....	95
<b>Table 4.3</b> The entries of a confusion matrix for binary classifications .....	98
<b>Table 4.4</b> Summary of models’ accuracy compared to baseline models .....	100
<b>Table 4.5</b> Performance of classification models .....	101
<b>Table 4.6</b> Performance of the Twitter’s identity model .....	103
<b>Table 4.7</b> Attribute values of the Twitter’s identity model .....	105
<b>Table 4.8</b> Performance of the Twitter’s interaction model .....	106
<b>Table 4.9</b> Attribute values of the Twitter’s interaction model .....	108
<b>Table 4.10</b> Performance of the Twitter’s visibility model .....	109
<b>Table 4.11</b> Attribute values of the Twitter’s visibility model .....	110
<b>Table 4.12</b> Performance of the Instagram’s identity model .....	112
<b>Table 4.13</b> Attribute values of the Instagram’s identity model .....	113
<b>Table 4.14</b> Performance of the Instagram’s interaction model .....	114
<b>Table 4.15</b> Attribute values of the Instagram’s interaction model .....	116
<b>Table 4.16</b> Performance of the Instagram’s visibility model .....	117
<b>Table 4.17</b> Attribute values of the Instagram’s visibility model .....	118
<b>Table 4.18</b> Distributions of training data and predictions .....	120

<b>Table 4.19</b> Correlation matrices of richness scores .....	121
<b>Table 4.20</b> Proportions of posts categorized as rich in each affordance dimension .....	122
<b>Table 5.1</b> Statistical tests on the differences of scores by platforms.....	129
<b>Table 5.2</b> Statistical tests on differences of scores by the celebrity types.....	131
<b>Table 5.3</b> Linear regression model of the engagement scores and richness scores.....	135
<b>Table 5.4</b> Linear regression model of the engagement scores and richness scores.....	138
<b>Table 5.5</b> Linear regression model of the engagement scores and richness scores.....	141
<b>Table 5.6</b> Linear regression model of the engagement scores and richness scores.....	144
<b>Table 5.7</b> Summary of models' performance .....	147
<b>Table 5.8</b> Linear regression model of mean engagement scores and variances of scores ....	150
<b>Table 5.9</b> Linear regression model of mean engagement scores and variances of scores ....	153
<b>Table 5.10</b> Linear regression model of mean engagement scores and variances of scores ...	156
<b>Table 5.11</b> Linear regression model of mean engagement scores and variances of scores ...	159
<b>Table 5.12</b> Summary of models' performance .....	161
<b>Table 5.13</b> Linear regression model of the changes in followers and mean richness scores .	164
<b>Table 5.14</b> Linear regression model of the changes in followers and mean richness scores .	167
<b>Table 5.15</b> Summary of models' performance .....	169
<b>Table 5.16</b> Linear regression model of the changes in followers and variances of scores ...	172
<b>Table 5.17</b> Linear regression model of the changes in followers and variances of scores ...	175
<b>Table 5.18</b> Summary of models' performance .....	177
<b>Table 5.19</b> Summary of research questions, methods, results and proposed statements .....	180

<b>Table 6.1</b> Axial codes and descriptions .....	191
<b>Table 6.2</b> Coding process .....	195

## LIST OF EQUATIONS

<b>Equation 2.1</b> Facebook’s official engagement formulas.....	65
<b>Equation 4.1</b> A linear combination of weights and the uses of communicative artifacts...	85
<b>Equation 4.2</b> Facebook’s official engagement formulas.....	124
<b>Equation 4.3</b> Instagram’s engagement formulas.....	124
<b>Equation 4.4</b> Twitter’s engagement formulas.....	124

# CHAPTER 1

## INTRODUCTION

This document provides details of a research study of microcelebrity practices on multiple social media platforms. Microcelebrity is a set of self-presentation techniques using technologies like social media sites to gain and/or maintain an audience (Senft, 2008). It is now common for celebrities to bypass mass media and take control over their promotional discourse. As evidence of this, numerous celebrities have emerged within, and as the results of their actions on, social media (Burgess & Green, 2009; Kassing & Sanderson, 2010; Snickars & Vonderau, 2009).

This work adopts an explanatory sequential mixed-methods research design to assist the interpretations and explanations of quantitative results with qualitative studies (Creswell, 2013). I begin with the design and development of a richness framework by borrowing concepts from the Affordances Theory (Gibson, 2014a), follow with a series of statistical analyses to examine the relationships between celebrities' social media uses and outcome of microcelebrity practices. The causality of the relationships is explained by interviews with audience members or fans.

The first section of this chapter presents an overview of microcelebrity practices on social media and the motivation behind the study, while articulating significant gaps in the literature. Second, I present an overall picture of the research problem and introduce the key inquiry of the study. In the third section, I present a methodological model including a richness framework as a primary tool for the quantitative component of this study, and the design of the follow-up

qualitative component. This section also includes an overview of the highlighted findings. The fourth section explains the key terminologies that will be used throughout this document. Then, I present the relevance of this work, including the expected theoretical and practical contributions. The last section presents an organization of the document.

## 1.1 Background and Motivation

Web 2.0 innovations, and particularly social media, have introduced a contemporary shift from broadcast to participatory media, through which content can now be produced, manipulated and distributed by everyone with Internet access (Jenkins, 2006). As media change, so does the celebrity culture (Marwick, 2015b). The shift has brought about major changes to the celebrity management model as celebrities can bypass attention brokers through the practice of *microcelebrity*.

The term was first coined by Theresa Senft (2008) from her study on *Camgirls* – a group of female personalities broadcasting personal webcam over the Web to the general public. The theory of Microcelebrity is a set of practices in which people construct their public persona as a *commodity sign* or product to be consumed by others (Hearn, 2008), using strategic intimacy to appeal to followers (Senft, 2008), and regarding their audience as fans (Marwick & boyd, 2011). Although Senft's work directly investigated *ordinary people* gaining status online, microcelebrity can be practiced across the spectrum of fame (Marwick & boyd, 2011). Traditional celebrities or those who have benefited from mainstream media attention are increasingly using social media sites for their promotional discourses. Social media sites have



enabled celebrities of any form to take more control over the presentation of their persona and the relationship they have with fans (Turner, 2013).

It is important to note that microcelebrity practices have placed *celebrity status* on a continuum from globally famous down to local/niche celebrity, rather than as a binary quality i.e., you are or you are not a celebrity (Marwick & boyd, 2011). The literature collectively suggests three core microcelebrity practices: identity construction, interaction with fans, and visibility promotion (Abidin, 2016; Marwick, 2013; Mavroudis & Milne, 2016; Page, 2012; Usher, 2015). Identity work on social media can be seen as a social act of positioning the self in relation to others (Page, 2012). The most direct form of identity construction is through sharing information (Marwick, 2015b) that reflects one's identity, or what they want others to believe reflects their identity (Mavroudis & Milne, 2016). Interaction develops and maintains audience through responding or reaching out to fans. Most social media sites provide conversational mechanisms, allowing users to interact or start a conversation around original content e.g., @mention and reply. Visibility enables microcelebrity persons to be found by others beyond the existing fan base in order to expand audience. With social media, celebrities can compete for visibility by engaging in the acts that promote public exposure beyond their followers – for example, by using hashtags or participating in online communities.

Previous studies, however, show that not all practitioners necessarily engage in all three core practices to the same degree, at least not on the platform of the study. For example, some studies found celebrities rarely interacted with fans on Instagram (Marwick, 2015a; Ward, 2016), but others have documented their interaction work on Twitter (Huba, 2013; Pegoraro, 2010). Moreover, it has become more common that people use multiple platforms simultaneously. Pew reported 66% of Twitter users also used Instagram (Greenwood, Perrin, & Duggan, 2016).

Microcelebrity practitioners have adopted this trend too (Marwick, 2015a). Using multiple platforms gives celebrities an opportunity to expand their audience, while overcoming the limitations of a particular platform. In other words, some platforms may support some practices better than others. This points to a gap in literature and calls for a cross-platform study.

As such, this dissertation is designed to be a cross-platform study to better reflect the roles of social media on microcelebrity practices in the broad media landscape. Amongst many social media platforms, this study puts emphasis on Twitter and Instagram. While these two platforms share similar characteristic as micro-content service (i.e., short text vs. still image), they differ by the focus/nature of the platform (i.e., textually vs. visually driven).

## 1.2 Research Problem

Microcelebrity practitioners have embraced multiple social media platforms, each of which differs by nature, functionality and users. Little is known about the roles of social media services and their communicative affordances in the outcomes of microcelebrity practices – *celebrity status*. In this research, I conceptualize celebrity status as the responses from audience and rely on two proxies: audience growth and audience engagement. The general question behind this inquiry is: *How do celebrities use social media for growing and maintaining audiences?*

The challenge in answering this question is the lack of a systematic way to examine social media activities that supports an analysis that looks beyond any specific platform. It is also important that the examination of practices must preserve different dimensions/aspects that the practitioners might engage in. To tackle this problem, I developed a richness framework to assess

social media uses from the affordances, or action possibilities, perspective (Gibson, 2014a). The development of the framework is part of the quantitative phase of my study, whose results inform the design of the follow-up qualitative analysis. Details of the research methods are presented in the following section.

### 1.3 Research Methods and Findings

As noted earlier, I employ an explanatory sequential mixed-methods research design to assist the interpretations and explanations of quantitative results with qualitative studies (Creswell, 2013). This study has been reviewed and exempted by the Syracuse Institutional Review Board (IRB#17-323). An electronic copy of the IRB authorization is included in the Appendix.

I begin with the design and development of a richness framework by borrowing concepts from Affordances Theory (Gibson, 2014a), which provides a foundation for assessing an information environment from the action possibilities perspective. Specifically, affordances are a range of action possibilities that an environment (e.g., Twitter and Instagram) allows users to perform by using *technological artifacts* (e.g., @mention and an ability to post pictures) (Kaptelinin & Nardi, 2012). For example, by providing an @mention artifact, Twitter affords an interaction affordance.

In this study, I also propose a notion of *affordance richness*. Similar to the concept of *media richness* – a measure of the richness of information carried by a communication medium (Daft & Lengel, 1986), *affordance richness* measures the richness of affordances made possible by a medium. For example, when a celebrity creates a tweet (a medium) with an @mention (an

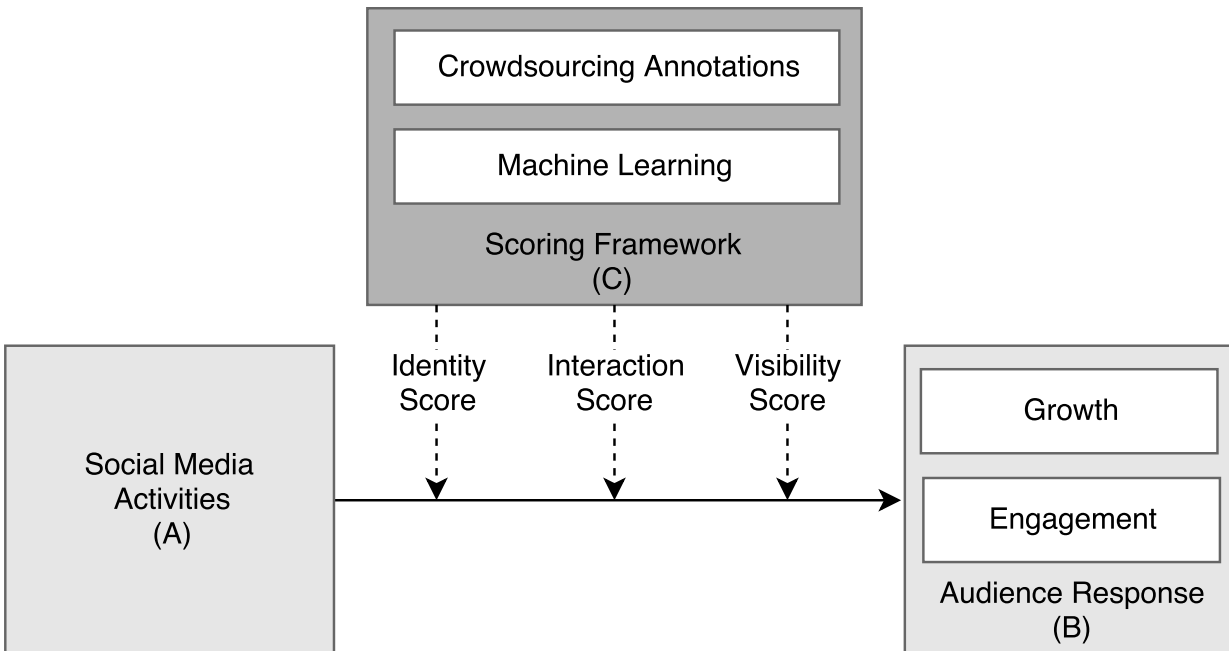
artifact) to interact (an affordance) with someone in the audience, I would say that the tweet is rich in interaction affordance. In this example, I would be measuring the richness of the interaction affordance. Simultaneously, I could also be measuring the richness of other affordances.

In the context of microcelebrity, I suggest social media sites providing three affordances mapped to the core microcelebrity practices: identity construction, interaction, and visibility promotion. The theoretical lens of Affordances allows the analyses to capture the ability of social media to serve the information needs of celebrities, while preserving different dimensions of practices along which the celebrities might engage. Each social media platform provides a different set of technology artifacts that contribute to affordances in different ways (Fayard & Weeks, 2014; Kaptelinin & Nardi, 2012; Zheng & Yu, 2016). Within each platform, users can construct a post (e.g., tweet or Instagram post) in different ways by using different combinations of technology artifacts (Brinker, Gastil, & Richards, 2015). Each post varies in affordance richness along the three dimensions: identity, interaction, and visibility affordances.

With the three dimensions of affordances in place, the framework then organizes the technology artifacts of social media into groupings, mapped to the three affordance dimensions. The affordance-artifact groupings are based on the HCI literature concerning the technology artifacts of social media (boyd, Golder, & Lotan, 2010; Honey & Herring, 2009; Hu, Manikonda, & Kambhampati, 2014; Kwak, Chun, & Moon, 2011). For example, @mention and @reply are organized to the grouping associated with the interaction affordance. The affordance richness measures were predicted by machine learning classification models, attributed by the uses of technology artifacts within the associated affordance-artifact grouping. The models were trained with training data annotated by crowdsourcing workers.

Building on three theoretical concepts (i.e., Microcelebrity, Affordances and Media Richness Theory), I argue that the performance of microcelebrity is co-constructed by celebrities and their fans, and the performance is mediated by social media. On the one hand, celebrities utilize multiple social media to manage *parasocial relationships* (i.e., one-sided relationships) by appropriating affordances with different richness. For example, a celebrity might appropriate more richness in identity affordance on Instagram than on Twitter, or having more richness in interaction than visibility affordance. On the other hand, fans provide feedback by responding differently to different microcelebrity strategies. The fans' responses form a feedback loop which then shapes how a celebrity performs on social media. Together, this suggests that social media have gradually given the fans more control over the celebrity-fans relationships and moved the one-sided relationships a little closer to two-sided relationships, or at least an illusion of such.

The overall picture of this research is illustrated in Figure 1.1. The two boxes (in light grey) represent the relationship between microcelebrity practices (A) and responses from the audience (B). The richness scoring framework untangles the relationship by quantifying the activities with affordance richness scores in the three dimensions.



**Figure 1.1** Overall Research Scheme.

With this framework, I answered a wide range of questions regarding the uses of social media in the context of microcelebrity practices through the analyses of affordance richness scores. The dataset is a collection of Twitter and Instagram data from 33 mainstream and 45 Internet celebrities. The analytical methods included statistical tests and regression analyses.

The findings from the quantitative analyses informed the design of the follow-up qualitative study. The qualitative phase of the study was designed to assist the interpretation and provide the causality of the relationships through interviews with audience members. In total, I conducted 15 one-on-one interviews, each of which was roughly an hour long. The interviews were semi-structured and guided by a set of open-ended questions and follow-up questions to draw out more information from informants. All interviews were audio-recorded and subsequently transcribed by myself for further analysis. The coding process employed an approach that gradually allowed themes to emerge as realized through information reduction,

conceptualization, elaboration, and relating (Strauss & Corbin, 1990). My final codebook is based on the primary theme of the process of co-constructing microcelebrity performance by celebrities and fans, as mediated by social media. The codebook consists of three main codes, each of which comprises six sub-codes.

The results show that celebrities need multiple social media platforms to perform microcelebrity and manage parasocial relationships. Twitter is more suitable for some practices while Instagram is more suitable for others as reflected by celebrities using them more often. My findings also reveal differences between the practices of mainstream and of Internet celebrities. The analysis of audience engagement shows that audiences were more likely to engage with the posts categorized as rich in some affordance dimensions but less likely to engage with the posts categorized as rich in others. That is, the performance of microcelebrity will be more effective when celebrities using the right richness for the tasks, as judged by the fans. However, decisions to follow or unfollow accounts were independent of the richness of the posts, but affected by the consistency of the richness over time. Together, this reflects the nature of the mediated microcelebrity performance that affords the fans with more access to celebrities and gradually brings the parasocial relationships closer to two-way relationships through the feedback channels.

#### **1.4 Terminologies**

For the rest of the document, I will use the terms *Internet famous* or *Internet celebrity* to refer to those who started out as ordinary people, and became celebrities as a result of their activities on the Internet. The terms *mainstream famous* or *mainstream celebrity* describe people such as pop stars and actors who have benefited from traditional mainstream media. I will use *celebrities* to refer to both types of celebrities.

In the context of my study, I will use the term *affordances* to refer to the abstract high-level action possibilities that social media afford for celebrities to perform. More specifically, the affordances comprise three dimensions, mapped to the core microcelebrity practices: identity, interaction, and visibility, each of which is enabled by *technical artifacts* of social media e.g., @mentions and embedded content. The term *affordance richness* means the richness of a post in an affordance dimension – for example, *identity richness* of a tweet means the richness in identity affordance of the tweet, or how rich the tweet is in its use of identity affordance.

## 1.5 Contributions

The motivation and anticipated contributions for this study comprise both theoretical and practical issues. The theory of microcelebrity has been previously studied within the limited space of a specific platform (Abidin, 2014; Marwick & boyd, 2011; Ward, 2016). Cross-platform studies would tackle some unanswered questions. For example, literature shows that not all practitioners use the same mix core practices to the same degree (Marwick, 2015b; Rahmawan, 2013) but it remained unclear whether or not the findings were limited by the platform-specific nature of the studies. Also, among the three core practices, little is known about whether or not any particular practices are more important than others on any particular platform. Cross-platform studies like this work shed light on microcelebrity practices in the broad media landscape by allowing for comparisons of their practices on different platforms. Specifically, the richness framework answered the questions by revealing different usage patterns on Twitter vs. Instagram, whether the audience responded to different strategies similarly or differently, and how the audience responds to changes in strategies. For example, we learned that microcelebrity



strategies had no effects on audience's decisions to follow or unfollow the accounts, but that the consistency of strategies did play an important role in their decisions.

Practically, this study provides a broad insight into microcelebrity practices from both practitioner's and audience's perspectives. As Marwick and boyd (2011) note, microcelebrity practices are *learned techniques*; such knowledge could be useful for generating recommendations on best practices. The results from this study are also useful for anyone trying to build and engage with a larger audience, for example, politicians, activists, scientists, and even startup companies.

This work also makes theoretical contributions to a growing body of literature around the theory of Affordances with the development of the notion of *affordance richness*. Although the theory provides a useful foundation to assess information environments, it does not provide a systematic way to examine how the affordances, when undertaken, enable users to engage in social actions. As such, I adopted the notion of richness from the theory of Media Richness and defined *affordance richness* as the ability of a post to deliver the information necessary in affording a particular action by using some artifact. Depending on the way it is constructed, a post might be rich in identity affordance, for example, or in other dimensions of affordances.

Lastly, the richness framework makes a meaningful methodological contribution by offering a tool to study social media activities from the perspective of technological affordances. This framework provides a way for researchers to examine social media actors in different contexts, such as politicians, CEOs, activists, and non-celebrity users. For example, researchers can leverage the framework to examine how Russian troll accounts grew their network on Twitter during the 2016 presidential election. Online marketing can benefit from using the

framework to study Twitter/Instagram uses by the top brands and researchers studying social movements may find that successful activist groups use messages richer in some dimensions than others. Another fruitful area may be looking at how actors' emphasis on richness dimensions may change over time and whether or not such changes are predictive of changing markets or social conditions.

## **1.6 Document Organization**

The remainder of the document is organized as follow: Chapter 2 reviews literature concerning microcelebrity, social media studies, and Affordances and Media Richness Theory. It also summarizes the direction of the study as well as documenting specific research questions. Chapter 3 outlines the design of the research methodology and discusses an overview of each phase of the study: quantitative and qualitative. The quantitative methods are presented in Chapter 4, including the design and development of the richness framework as well as the richness scores analyses. Chapter 5 presents the results from the quantitative analyses and summarizes the main findings, which are used to inform the design of the follow-up qualitative study, outlined in Chapter 6. The results from the qualitative analysis are presented in Chapter 7, where I also discuss how they support and contradict the results from the prior quantitative analyses. Chapter 8 presents the discussions around the methods and findings of this dissertation. The document ends with a conclusion of this dissertation research.

## **CHAPTER 2**

### **LITERATURE REVIEW**

This chapter lays the foundation for the theoretical direction of the study through the review of three bodies of literature. First, I introduce the theory of Microcelebrity, a set of self-presentation techniques using technology like social media sites, and a review of related work. Then, I present theoretical models drawing on a conceptual lens of the theory of Affordances and touch on the theory of Media Richness. The last section presents a collection of social media studies in three related areas: a) collection of technological artifacts as part of the perspective adopted to study social media; b) followership and, c) follower engagement as the proxies of microcelebrity status. The chapter ends with an introduction to my research questions.

#### **2.1 Celebrity Studies**

Celebrity culture has always been linked to the media industry. They have traditionally been the product of promotions and publicity driven by mass media industries (Turner, 2004). The moment one becomes a celebrity is the “point at which media interest in their activities is transferred from reporting on their public role ... to investigating the details of their private lives” (Turner, 2013, p. 8). In his study of celebrity culture, Marshall (2006) argues that the invasive lenses of mainstream media provide the public with the chance to see what celebrities are truly like outside of their constructed world. One way to control their media persona in the face of media invasiveness is by employing the layers of representation e.g., agents, managers,

and publicists, who present a carefully *constructed personality* to the media, and thus the public (Turner, 2013). Indeed, *celebrity management* is a highly controlled and regulated institutional model (Marwick & boyd, 2011).

The rise of social networking sites has introduced the contemporary shift from broadcast to participatory media by means of which the content can be produced, manipulated, and distributed by the public (Bruns & Burgess, 2011; Jenkins, 2006; Page, 2012). As media changes, so does celebrity culture (Marwick, 2015b). Celebrification, or the process of turning oneself into a celebrity (Driessens, 2013), is no longer solely related to mass media, but also now reflects a more diverse media landscape. The transition in celebrity culture is what Gamson regards as *democratization in celebrification* (Gamson, 2011) and what Graeme Turner refers to as the *demotic turn* (Turner, 2010). They suggest that this emerging promotional culture has resulted in an increasing number of unexceptional people becoming famous, and stars who have been made ordinary. As evidence of this, numerous celebrities have emerged within, and as result of their actions on, social media (Burgess & Green, 2009; Kassing & Sanderson, 2010; Snickars & Vonderau, 2009) through *microcelebrity practices*.

### **2.1.1 Theory of Microcelebrity**

Microcelebrity is a set of self-presentation techniques, first coined by Theresa Senft (2008) as “a new style of online performance that involves people *amping up* their popularity over the web using technologies like video, blogs and social networking sites” (p. 25). It is a set of practices in which people construct their public persona as a *commodity sign*, or product, to be consumed by others (Hearn, 2008), use strategic intimacy to appeal to followers (Senft, 2008),

and regard their audience as fans (Marwick & boyd, 2011). The practices are typically engaged through the broadcast of “a continuum of selves” (Raun, 2018, p. 106), and success is “measured in likes, shares, follows, comments and so on” (Cottom, 2015, p. 2). Although Senft’s work directly investigates ordinary people gaining status online, traditional celebrities, or those who have benefited from mainstream media attention, are increasingly using social media sites for their promotional discourses and give audience access to their everyday lives (Burgess, Mitchell, & Münch, 2018). It is now common for celebrities of all types to bypass the mainstream media, and interact and communicate with the public directly. As a result, they have more control over the presentation of their persona and the relationship they have with fans (Turner, 2013).

Following Marwick and boyd (2011), I think of *celebrity* as a practice. It is what a person does rather than what a person is, and *celebrity status* exists as a continuum between globally famous down to a local/niche celebrity, rather than a binary quality (i.e., you are or you are not a celebrity). In the age of social media, everyone with Internet access can engage in such practices and become a microcelebrity practitioner. But only those who successfully present a consumable version of self gain status (Gamson, 1994). Notably, what considered as consumable varies by the social context within which the practitioners operate, but typically involves self-promotion through carefully constructed personas. This could take different forms, e.g., textual, visual, or video, depending on the technological affordances of the technology they employ, each of which allows users to perform certain actions (Norman, 2013).

The sign of celebrity status also varies by platforms and norms of platform’s members. For Camgirls (i.e., female personalities who broadcasted themselves on the Web), it is web viewership, or the number of unique visitors, that indicated their popularity (Senft, 2008). On Tumblr, a publicly visible status measure is number of likes on a post (Marwick, 2015b). For

other social media platforms that do not employ mutual relationship-based dynamic like Twitter and Instagram, the numbers of followers are usually regarded as a sign of status and become stand-ins for social status, signaling to the public that the users are worthwhile (R. Li, 2018; Marwick, 2013). Most sites also offer a mechanism for the audience to react to the posts such as *likes* and *retweets*, which function as a form of social feedback (Bakhshi, Shamma, & Gilbert, 2014).

Amongst many social media platforms, previous research highlights the roles of Twitter and Instagram as an important venue for developing parasocial relationships between celebrities and their fans (Ward, 2016). When it comes to Twitter, fans are given an access to a celebrity's personal life, directly interact with them, and believe themselves to be a part of the network. Being a real-time updater, Twitter helps create the sense of *being there* with celebrities and becomes an intimate form of communication for celebrity-fan relationships (Stever & Lawson, 2013). Twitter is also a space where the blurring of personal and professional roles is encouraged and rewarded (Gregory, 2018). Instagram, an image-driven platform, takes this relationship a step further by providing an actual look into celebrity's lives in addition to a textual update (Marwick, 2013; Ward, 2016). Alice Marwick (2015a) highlights the importance of studying Instagram as a means to move away from the focus of online identity as *written into being*. She argues that the Internet is becoming a visual medium and that an increasing number of people tend to express themselves through images rather than textual updates. However, the downside of Instagram, in the context of microcelebrity, is the limited opportunity for audience interaction.

Scholarship suggests that celebrities who develop reputations by performing themselves (e.g., pop music and sports stars) articulate their public persona with discourses of authenticity, or the expression of what they truly are in order to give the public an impression of *insider*

(Gamson, 1994). Authenticity, however, is not necessary a *property* of performers but could be a role developed for the performance (Usher, 2015). The closer the constructed personality and private self are together, the better (Marshall, 2006; Turner, 2013). This is also true for Internet famous practicing microcelebrity, as they turn themselves into media content to be consumed by the audience.

Consistent with Goffman's (1959) presentation of self, people continuously maintain the impressions they foster throughout the performance by maintaining a consistent identity. That is, fans expect high levels of *authenticity*, or at least *theatrical authenticity*, from their celebrities, be they traditionally famous or Internet famous. For example, fans of fashion bloggers reported authenticity as a value that differentiates bloggers from fashion magazines, with affordable goods on *average women* (Marwick, 2013) whom they can be more related to (Djafarova & Rushworth, 2017).

It has become common for celebrities practicing microcelebrity to use multiple social media platforms as a means to amp-up the fame they have achieved elsewhere (Marwick, 2015a), such that the audience on one platform overlaps with the audience on others. In an online world, public personas are utterly integrated, as it is almost impossible to compartmentalize different parts of one's online self (Senft, 2008). That means that celebrities need to articulate a consistent and authentic public persona across multiple platforms simultaneously (Marshall, 2006; Turner, 2013).

### 2.1.2 Core Practices of Microcelebrity

In addition to the need to maintain consistent and authentic public personas, scholars suggested the core properties of accruing celebrity status are *attention seeking* and *visibility promotion* (Marwick, 2013; Page, 2012) through ongoing fan management, self-presentation, and constant promotions. Social media are an arena of public attention, but attention itself is a scarce resource as it gets distributed and draws on various competing issues. Certain pieces of information have to compete with others to become visible (Brighenti, 2010).

Attention can be acquired through the *interaction* which treats the audience as an aggregated fan base to be developed and maintained, and the construction of *identity*. Interaction develops and maintains audience through responding, or reaching out to fans. Most social media sites provide conversational mechanisms, allowing users to interact or start a conversation around original content e.g., @mention and @reply. Additionally, the uses of second person pronouns (e.g., you and guys), asking questions, and asking for feedback or opinions to display the inclusiveness and create a sense of conversation between celebrities and audiences (Raun, 2018).

The identity work on social media can be regarded as a social act of positioning the self in relation to others (Page, 2012). The most direct form of identity construction is perhaps through sharing information (Khamis, Ang, & Welling, 2017; Marwick, 2015b), which reflects one's identity, or what they want others to have impression about them. Hackley et al. (2017) suggested celebrity selfies (i.e., the picture of oneself taken by oneself) represent a performance of mediated identity where celebrities use their lives as "the dramatic material" (Hackley et al., 2017, p. 51).



Visibility is to enable the identity and interaction work to be found by the larger public. If we think of such work as a piece of information, *visibility work* is to promote and compete for public attention. With the demotic turn (Turner, 2010) in our media culture, everyone with Internet access can compete for visibility, to varying degrees, by engaging in the acts that promote public exposure. Most social media platforms support a mechanism to increase visibility. On Twitter and Instagram, the hashtag is a wide spread convention to connect posts, highlight a common theme, make the posts appear in the search feature, and thus promote the visibility of the posts and authors beyond the existing fan base (Page, 2012).

Indeed, these practices are undertaken by both mainstream and Internet famous. While they help elevate ordinary persons to achieve celebrity status, they also bring the stars closer to fans by revealing their *ordinary people* aspect as a means to promote their authenticity (Turner, 2013). Examples include Camgirls broadcasting personal webcam on the Web (Senft, 2008), *Instafame* as a means to gain status on Instagram (Abidin, 2016; Mavroudis & Milne, 2016), uses of YouTube by amateur and professional performers (Burgess & Green, 2009; Marwick, 2015b) and mainstream celebrities on Twitter (Marwick & boyd, 2011; Seidel, Berente, Debortoli, & Srinivasan, 2016).

In the next section, I present a review of microcelebrity studies particularly on Instagram and Twitter. These two platforms are important venues for developing parasocial relationships of celebrities and their fans (Ward, 2016). They are a good point of comparison due to the differences in nature, meaning Twitter is a textually driven and Instagram is visually driven. Yet, they share some similarities. Both are relatively *micro-content* service. While Twitter is a

textually driven platform with 140-character limit<sup>1</sup>, Instagram is an image, or two-minute video sharing platform (thus *micro* when compared to other content community sites like YouTube (Kaplan & Haenlein, 2010)). Both do not employ a mutual relationship-based dynamic. For example, user A can follow user B, but user B is not required to follow user A. On the news feed, users see posts from a set of users they elect to follow. Also, both increasingly gain users every day, and as a result have become a great venue for people seeking an audience.

### 2.1.3 Related Studies

Theresa Senft coined the term *microcelebrity* from her study of *Camgirls*, a group of popular web personalities who broadcasted themselves on the web to the general public (Senft, 2008). She suggested that *Camgirls* consistently described themselves in no way similar to a film or television star, yet they do not present themselves as an ordinary person either. They also considered themselves *more real* than television personalities. Although there exists a vast quantity of studies looking at how people use social media for their promotional discourses, relatively little studies have actually adopted the term *microcelebrity*.

With social media, one can construct their identity in the way they wish others to have impression about them, they can build trust, rapport, and relationships with members of the public without being mediated by the mainstream media. In the following sections, I present related work examining microcelebrity practices on Instagram and Twitter by both mainstream and Internet famous persons.

---

<sup>1</sup> As of the time of this study, Twitter had a 140-characters limit but it was changed to 280-characters in November 2017.

### *A. Microcelebrity on Instagram*

Instagram is a photo sharing platform with around 600 million monthly active users around the world. It offers a great opportunity to achieve *Instaframe*, defined by Marwick as “the condition of having a relatively great number of followers on the app” (Marwick, 2015a, p. 137). Although the top Instagram users are mostly mainstream famous people (e.g., pop stars), an increasing number of ordinary people turn to Instagram in an attempt to gain audiences and become famous (i.e., Internet celebrity). Mavroudis and Milne (2016) conducted interviews with Internet celebrities in Los Angeles, who suggested that managing and maintaining their status was a type of immaterial labor. The practice typically involves promoting a sense of self and identity construction. One of the authors, who identified himself as an Internet celebrity, claimed that his content was a carefully crafted identity specifically designed for consumption by the public, with the ultimate aim of maintaining his status. The interviewees also reported the importance of maintaining a consistent identity – for example, they chose not to post some things, as those would not suit their profile, even though they wanted to. Crystal Abidin (2014) conducted a case study and found that her subject chose not to publish her clubbing photos in order to maintain her constructed persona as a role model for under 18 followers. She also found that Internet celebrities tended to form an exclusive network as they only posted photos with fellow celebrities, rather than friends who were not well-known for the Instagram followers.

Studying the microcelebrity practices from the marketing perspective, Abidin (2015) observed an attempt to naturalize the advertorials of Internet celebrities by referencing their children and family. Through discourse analysis, she found a strategic use of hashtags as a way

to mask the distinction between personal and commercial material. Specifically, they tended to use a particular hashtag in personal, non-sponsored images about their mundane activities with children, and later on transplanted the hashtag to their commercial posts. Marwick (2015a) conducted case studies on three highly followed Instagram accounts (greater than 10k followers) and noted that microcelebrity practices on Instagram took the forms of creating personae, sharing personal information through photographs, and strategically appealing to the audience.

Although previous work suggested that one of the key practices of microcelebrity was interaction with an audience (Marwick & boyd, 2011; Senft, 2008), Marwick (2015a) noted that not all practitioners engage in all the core practices to the same degree as evidenced by one of her case studies. She considered the lack of fan interactions as the dissimulation of their celebrity status. The lack of fan interactions is also documented by a study of traditional celebrities who tend to ignore their followers. Janabeth Ward (2016) content analyzed Instagram posts of American singers, Taylor Swift, Selena Gomez, and Ariana Grande. The study found different Instagramming styles by these individuals. For example, the most frequent topic for Taylor Swift is associating herself with other celebrities, while personal content is the most frequent topic for Ariana Grande. Amongst the three, fan interaction is consistently the least frequent topic. This work also found that amongst other categories, personal content posts received the most likes from fans. Another case study examined celebrity selfies of Medina, a Danish pop singer and songwriter (Jerslev & Mortensen, 2016). The authors suggested that Instagram selfies functioned as a successive documentation of celebrity's everyday lives. They keep fans updated and connected by creating a sense of intimacy, offering access to celebrity's lives while maintaining authenticity.

Another group of celebrities is political figures, who are "being sold in a political

advertisement” (Centeno, 2010, p. 72). The goal of political campaigns is to increase a candidate’s exposure to the public through the use of both mass media and, more recently, social media by engaging in microcelebrity practices. Politicians also make use of Instagram, although not as popular as Twitter. A study of impression management on Instagram by Singaporean politicians suggests that some politicians tend to post images about their private lives, while others restrict the content to the public professional lives only (Jung, Tay, Hong, Ho, & Goh, 2017). The authors conducted an experiment with 120 undergraduate students to measure the effects of Instagramming strategies on voting intention and impressions about the politicians. The results indicate that posting about public lives and attempts to interact with the public give a more positive impression about the politician’s character. However, they could not detect different effects of the strategies on voting intention. This study highlights the importance of maintaining a consistent identity across platforms. That is, politicians are usually advertised as a professional individual on mainstream media. As such, Instagram followers would expect to see a consistent persona on social media.

### *B. Microcelebrity on Twitter*

Twitter is a microblogging service that allows users to broadcast 140-character messages (tweets) to groups of other users who subscribe to their accounts (followers). There are about 330 millions monthly active users around the world (Twitter, Inc., 2017) which makes this social media site an excellent platform for anyone who wants to gain audience or seeks fame.

Detta Rahmawan (2013) examined Internet celebrities in Indonesia and suggested that the main uses of Twitter is to update the audience about their daily lives and position themselves by

articulating opinions. The author highlighted the roles of @replies and retweets to fans as a status symbol within the fan community, although they were not frequently used. Ruth Page (2012) examined hashtag uses by three groups of users: corporations, mainstream celebrities and ordinary people practicing microcelebrity. She found that ordinary people tended to use hashtags to position themselves through the commentaries around national events e.g., political elections, sporting events and television shows. However, when traditional celebrities employed such hashtags, they tended to project their identity in relation to commodities e.g., their performances, products and campaigns, as a part of their professional, front stage identities.

Looking at traditionally famous people, Marwick and boyd (2011) suggest that Twitter creates the celebrity-fan relationship with a sense of intimacy. The platform allows them to design the way they want to be perceived by fans. For instance, celebrities use @reply as a form of public acknowledgement to give back to loyal followers and create a sense of intimacy rather than appear to be uncaring or unavailable. Sun Jung (2011) proposed that Twitter has greatly supported the interaction between stars and their fans, especially those with different national or linguistic backgrounds. He found the most common types of tweets to be daily updates and direct communication with fans. Similarly, Bennett (2014) and Huba (2013) specifically study Lady Gaga, an American recording artist. These studies suggest that she skillfully utilizes Twitter by not only combining public/private elements of herself but also by maintaining direct and constant communication with her fans. Through the uses of hashtags, Marlee Matlin, a deaf American actor, started a campaign to raise awareness of deaf culture, political equality and media access (Ellcessor, 2018). The author notes that the use of hashtags enables Matlin to reach her target audiences and society at large.

In sport communication, Ann Pegoraro (2010) suggests that Twitter is a powerful tool for developing athlete-fan relationships where athletes can share their stories as openly as they wish without third-party mediation. The work shows that the most common type of athlete's tweets is *direct tweet*. Direct tweet is a group of tweets with specified receiver(s) usually starting with @mention. Such tweets are helpful for building and maintaining relationships with the fans (for example, responding to fans' queries). Frederick et al. (2014) also found that the majority of athletes' tweets were intended to interact with the public by asking questions and talking about their personal lives to create a sense of intimacy. Their findings are consistent with Kassing and Sanderson's (2010) work whose results suggest that athletes use Twitter to support interactivity with their fans. The majority of their tweets provide commentary and opinions which cultivate insider perspectives for their fans. The aforementioned studies have a common finding, athletes primarily utilize Twitter for increasing fan attachment by consistently interacting with their fans, and athletes construct their public persona by revealing private elements of their lives.

The use of Twitter by politicians is now commonplace. At best, Twitter provides the public with an opportunity to directly interact with, and engage in, political discourse with both candidates and elected officials (Ausserhofer & Maireder, 2013). Enli and Skogerbo (2013) showed that candidates running for election utilized Twitter to increase their visibility and engage in continuous dialogue. Highlighting the perquisites of microcelebrity practices, Conway et al. (2013) found that the most active presidential candidates in the 2012 U.S. primary election were not from the major parties (i.e., Democrat and Republican) but from alternate parties, such as the Green Party or the Libertarian Party. Similarly, Christian Christensen (2013) suggested that while candidates from alternate parties typically suffered from limited support and resources, social media platforms like Twitter offered them opportunities to gain attention and

move towards the political front. Compared to mainstream candidates, these third-party candidates tend to create the highest tweet volumes during the debates.

Of course, a large number of tweets is not a measure of audience engagement Christensen (2013) showed that candidates who use hashtags in creative ways tended to have higher rates of audience engagement, as measured by how often they were retweeted. Graham et al. (2013) suggested that candidates of the minor Liberal Party in the U.K. utilized Twitter to promote themselves during the 2010 general election more than their major party counterparts did. In Australia, elected officials are generally *noisier* than the public in that they tend to broadcast more than interact with the audience (Grant, Moon, & Grant, 2010). Likewise, recent scholarships exploring the tweets of Members of Congress (Golbeck, Grimes, & Rogers, 2010; Hemphill, Otterbacher, & Shapiro, 2013) found that they used social media as a broadcast mechanism, rather than as a mechanism for interaction with constituents. Glassman et al. (2009) also explored the tweets from members of Congress and found that they used Twitter to construct their identity by taking stances on controversial issues such as expressing concern about a specific bill under consideration or a general policy issue. Grant et al. (2010) suggest that, compared to the public, politicians tended to use Twitter for broadcasting information more than engaging in conversations. Taken together, the literature suggests that politicians use Twitter primarily to increase their visibility and construct their identity by broadcasting information.

### **2.1.4 Summary**

Literature has shown different ways celebrities, be they mainstream famous or Internet



famous, use Twitter and Instagram for their promotional discourses. Although relatively little studies used the term *microcelebrity* specifically – for example, Li (2018) refers to such practices as *do-it-yourself celebrity plan*, the practices indeed fit in with the definition of microcelebrity as a set of self-presentation techniques to amp-up their popularity by using technology like social networking sites. The discussed literature shows that not all practitioners necessary engaged in all three core practices – identity, interaction, and visibility – to the same degree, at least not on the platform the studies took place.

It has become common that people use multiple platforms simultaneously. Pew reported 66% of Twitter users also used Instagram (Greenwood et al., 2016). Although adopting multiple platforms could be somewhat challenging for celebrities, as they need to maintain a consistent and authentic public persona, using multiple platforms would give them an opportunity to expand and strengthen the relationships with their audience (Khamis et al., 2017), and overcome the limitations of a platform. In this study, I examine microcelebrity practices on multiple platforms, Twitter and Instagram, and conceptualizing microcelebrity as the three core practices: identity construction, interaction, and visibility promotion.

Recall the overarching question of this study is *How do celebrities use social media to grow and maintain celebrity status?* The challenge in answering this question is the lack of a systematic way to examine social media uses from the user-centric perspective. In the context of microcelebrity, an examination of the practices should preserve different dimensions along which people might engage in i.e., identity construction, interacting with fans, and promoting visibility beyond the existing fan base. Literature shows that certain social media behaviors could have different effects when conducted by different actors (Araújo, Corrêa, da Silva, Prates, &

Meira, 2014; Xu, Huang, Kwak, & Contractor, 2013). As such, the designs of any assessment frameworks should be contextualized to the setting of the study.

To systematically examine social media uses in the context of microcelebrity, I draw on the theoretical models including two bodies of literature: Affordances and Media Richness theory. Specifically, the theory of Affordances (Gibson, 2014a) is adopted as a primary conceptual lens to study the uses of social media for microcelebrity practices. This study also borrows the notion of *richness* from Daft and Lengel's (1986) work to quantify the ability of a communicative medium to deliver information.

## **2.2 Theoretical Models**

This study primarily draws on the theory of Affordances (Gibson, 2014a, 2014b) to assess the information environment from an affordances, or action possibilities, perspective. Specifically, affordances are the relationships between technical objects (or technical functionalities/artifacts) and a user's interpretation. The theory, however, does not provide a systematic way to assess the ability of the affordances to help users achieve a desirable result. Therefore, I also borrow the notion of *richness* from Daft and Lengel's (1986) study.

In the following sub-sections, I present the theory of Affordances and its related studies. The theory provides a useful foundation to assess information environment without limiting an analysis to any particular platforms. Then, I present Media Richness Theory. As shown later, the theory has generated a substantial body of literature in many different areas (e.g., organizations, friendship development and social media uses), some of which have posed a challenge to the theory with contradictory evidences. Specifically, the technology has been rapidly developed

such that the differences between traditional and new media are too great (Carlson & Zmud, 1999; M. El-Shinnawy & Markus, 1997; Fulk & Boyd, 1991; Kinney & Watson, 1992; Kishi, 2008; Markus, 1994). However, the notion of *richness* – or the ability of a medium to deliver rich information which is varied by its supports of communicative artifacts, is still a good fit (Alan R Dennis & Kinney, 1998; Kishi, 2008). As such, I only adopt the notion of richness and define a new term *affordance richness*. Details are as follow.

### 2.2.1 Theory of Affordances

The term *affordance* was first coined by Gibson, an ecological psychologist, in 1979. He explains that affordances are action possibilities suggesting how objects could be used (Gibson, 2014a). They are independent of the needs or goals of the user i.e., the object always affords what it does even if a user's needs or goals have changed. Although the existence of affordances is independent, their interpretations are relational to users. The interpretations emerge from interactions between the object and user, and so the same object could be interpreted differently by different users. For example, a chair always affords seating but only perceivably to humans, not to fish. Although the theory was developed in the context of animals and the natural environment, Gibson notes that the theory is applicable for studying human beings and the cultural environment.

When introducing the concept of affordances to HCI, Norman (1988) further suggests that affordances are a combination of *perceived* and *actual* properties, all of which provide strong clues about their functionalities and determine how they could be used. For Norman, the perceived properties are similar to Gibsonian affordances i.e., referring to the perception of how

the objects should be used, whereas the actual properties are the fixed materiality of the objects. The combination of the two properties offers the *real* action possibilities for users (Norman, 2013). In addition to affordances being perceptions about what we can do or how we can act through the objects, another important function is to place constraints on what could not be done (Norman, 2013). Constraints can be either objective or subjective. Some rely upon the accepted cultural conventions even if they do not affect the physical or semantic operations of the object. As different cultures have different sets of acceptable actions, the same object might be perceived differently.

The key difference between Gibson's and Norman's notions is the existence of affordances and user's perception. For Gibson, affordances are independent of the perception, i.e., they exist even if users do not perceive them. He, however, notes that the essential aspect of affordances is not their existence but it is the extent to which they provide information for the actors to perceive. This is what he refers to as *direct perception*. Norman, on the other hand, argues affordances must be real and perceived otherwise they do not exist. Regardless of the difference, the essential element is perhaps the perception of affordances. People naturally establish affordances by developing a *mapping*, or a relationship between actions and results (Gibson, 2014a). We create the mapping by picking up information from the object itself and other users who typically provide the richest and most elaborate information, in other words, "behavior affords behavior" (Gibson, 2014b, p. 58).

In an attempt to untangle the relationships between affordances and perception, Gaver (1991) develops a framework for classifying affordances, re-defined as "properties of the world that are compatible with and relevant for people's interaction" (Gaver, 1991, p. 79). The framework consists of four elements: *perceptible affordances* refer to real and perceived

properties; *hidden affordances* refer to real but not perceived properties; *false affordances* mean incorrectly perceived properties as they do not exist; lastly, *correct rejection affordances* refer to properties correctly not perceived as they do not exist. He demonstrates that the concept of affordances is a useful theoretical lens for analyzing the user-centered design of computer mediated communication tools. He suggests that it provides a framework useful for guiding a design to focus not only on technology or users alone, but also on the interactions between them. McGrenere and Ho (2000) make a similar argument to separate affordances from the perceptual information but comment on Gaver's (1991) *false affordances* that the perceptions are not wrong; rather, it is the information that is wrong. In other words, false affordances should be re-defined as the misinterpretation of an object, which occurs when users pick up misinformation.

The following literature put an emphasis on *perceptible affordances* or the real and perceived properties of an object. On this basis, scholars have adopted the theory of Affordances to examine the process through which technology affords users the ability to perform communicative tasks. One of the widely used variations of the affordances concept is *social affordances* or the relational properties of the object that enable interactions amongst group members given their social characteristics (Bradner, Kellogg, & Erickson, 1999). Bradner and his colleagues note that it is important to distinguish social affordances from the original definitions to emphasize the social and cultural aspects of the appropriations (Bradner, 2001; Bradner et al., 1999). In other words, while Gibson's notion concerns the interaction between an object and user, Bradner and his colleagues (2001; 1999) are concerned with the interactions between users as afforded by the object. Similar to affordances varying by individuals, social affordances are bounded by the context and social norms. From their study on the adoptions of a chat tool, Bradner et al. (1999) suggest each subject group collectively develops understanding

and legitimacy, as members gain experience with the system. They note that a lens of social affordances offers a way to examine the interplay between technology properties, communicative practices and social characteristics of a group, particularly with respect to practices that the group recognizes as legitimate. Putting emphasis on the context as certain technologies are associated in the imagination of users, Nagy and Neff (2015) propose a concept of *imaged affordances*, which emerge between technology, users' goals and designers' intentions. Wellman and his colleagues (2001; 2003) study how social affordances in computer-supported interpersonal communication affect the ways in which people connect with each other. Wellman (2001) suggests that technology brings about greater bandwidth for non-face-to-face communication but it is how we appropriate the technology that creates and sustains community. Another study (Wellman et al., 2003) suggests that the Internet should be considered as a multi-dimensional medium which offers five social affordances for developing communities: broader bandwidth for rapid exchanges of large amount of data, personalization, wireless portability and connectivity in both time (always connected) and space (globalization). Hsieh (2012) proposes a framework for studying digital inequality particularly in the context of interaction. He suggests that more skilled users typically engage in more digital capital-enhancing activities, e.g., using social media to maintain social relationships. This study draws on the concept of *social affordances* and identifies an additional digital skill, namely, social networking skills as "the ability to use ICTs to facilitate social interactions" (Hsieh, 2012, p. 11). Although an ability to use digital media can certainly allow individuals to communicate and interact with others, but we cannot assume that the interactions between communication partners will be successful or sustained. To capture how the technology and society are related, Hutchby (2014) defines the concept of *communicative affordances* at the intersection of technological determinism and

social constructivism where affordances are both functional and relational. That is, they enable users to perform a range of possible actions but the range could differ between individuals. He agrees with Gibson that affordances are best observed in communication between users. This is manifested by his use of the term *communicative* to put an emphasis on the impact of technology for communication.

More recently, Kaptelinin and Nardi (2012) explain that affordances are mediated by cultural means and enabled by technology artifacts. Such artifacts are typically designed to support specific tasks/operations. However, users may not necessarily use them for, nor be limited by, the intended purpose (Markus & Silver, 2008) as unintended functionalities often arise after user engagement (O’Riordan, Feller, & Nagle, 2012). Social media sites are a great example to illustrate this point as a majority of their functionalities are emergent and shaped by user appropriation choices (O’Riordan et al., 2012).

The environments of social media have been rapidly evolved. Each site continuously improves its UI and back-end service, and so do user practices. Many features of social media sites have been progressively developed and integrated into their architecture over time by user conventions (Bruns & Burgess, 2011). Twitter, for example, has been gradually and culturally developed over time. For example, hashtag was originally proposed to the Twitter community by Chris Messina, a software developer, as a system of channel tags for “improving contextualization, content filtering and exploratory serendipity within Twitter” (Wikipedia contributors, 2016), and was integrated into Twitter’s architecture later.

Social media scholars adopt a theoretical lens of the Affordances to contextualize their studies in relation to higher-level patterns of behavior as opposed to the idiosyncratic features of

the sites (Ellison & Vitak, 2015; Fayard & Weeks, 2014). The lens enables researchers to capture relationships between the technological materiality and users while avoiding limiting themselves to a particular site and set of users at a particular moment in time. More importantly, social media have been rapidly increased in popularity and usage, and so do the relationships between technological materiality and users. On the one hand, the ways in which users expect certain algorithmic affordances affect how they approach these platforms. On the other hand, the feedback-loop characteristics of machine learning systems like Facebook make user beliefs an important component in shaping the overall system behavior, as end-user activity is generative of the system itself. That is, the affordances may not just affect how users approach social media platforms, but performativity also helps shape the platforms themselves (Nagy & Neff, 2015). Social media affordances have been gradually developed over time and shared amongst users. Some of them are common across platforms and others are exclusively available on a particular platform.

Work in this area uses the lens to understand the potential uses of technology by examining how social networking sites afford users the ability to perform communicative tasks. Specifically, instead of focusing on any particular technology, an affordance approach allows researchers to focus on the dynamics or types of communicative practices and social interactions afforded by the technology (boyd, 2010; Schrock, 2015; Treem & Leonardi, 2013).

Some studies use the term *affordances* almost synonymously with the technological features or technical features (e.g., Black, Mascaro, Gallagher, & Goggins, 2012; Gleason, 2013). Specifically, such features are the materiality or properties of the technology or a medium such as a button and hashtags. However, affordances are broader than the properties of technology such as “buttons, screens and operating systems” (Schrock, 2015, p. 1233); they are enhanced



and conditioned by the properties of technology (boyd, 2010). Other works in this view focus on the abstract high-level affordances enabled by technical features, or the kinds of communicative practices and user interactions enabled by the technological materiality. For example, boyd (2010) examines social media as a networked public and derives a common set of affordances of social media: *persistence* refers an ability to access a message after posting; *replicability* or an ability to duplicate content; *scalability* or an ability to make information visible for others; and *searchability* or an ability to locate information. In the context of organizational communication, Treem and Leonardi (2013) examine social media sites and suggest four common affordances: *visibility* and *persistence* which are similar to boyd's (2010); *editability* refers to an ability to craft a message before posting and edit after posting; and *association* or a connection between users and their content. In the blogosphere, Graves (2007) identifies three affordances of blogging, particularly for journalism: *reader input* or a fact-checking by the crowd; *fixity* which is similar to *persistence* (boyd, 2010; Treem & Leonardi, 2013); and *juxtaposition* which refers to an ability to put together several pieces of news to tease out implications.

Highlighting the relational property of affordances, Fayard and Weeks (2014) develop the concept of *affordances for practice* or the high-level technical properties of the sites which afford a specific user or group to perform goal-oriented actions within particular social, cultural and historical contexts. This perspective considers affordances as embedded in, and emerging from, social processes. Thus, it allows for a systematic examination of social media affordances within an associated context. To illustrate this point, Mansour et al. (2013) and Majchrzak et al. (2013) both examine Wikis, but with different emphases. The first study is interested in the affordances of Wikis for individual uses and so suggests four affordances: *commenting*, *accessibility*, *viewability* and *validation*. The other study, on the other hand, is more interested in

group work, and comes up with four different affordances: *meta-voicing*, *triggered attending*, *network-informed associating* and *generative role-taking*. These studies are a great evidence of affordances being relational and show that the examinations of affordances should be conducted within the associated context.

A study on Couple, a dating app, develops a notion of *vernacular affordances* to link the materiality of social media sites to the affordances derived from the user-centric perspective (McVeigh-Schultz & Baym, 2015). They suggest deriving the affordances from the sense-making processes of users. Moving in this direction, Ellison and Vitak (2015) study the processes of social capital through an examination of Facebook affordances, e.g., an exchange of informational and social support through the uses of three features: the profile, the friends list and the broadcasting updates. Bucher and Helmond (2017) examine the case of Twitter's *favorite* button. They suggest that the name and appearance are highly coupled with the perceived range of possibilities afforded by the feature. Particularly, in November 2015, Twitter changed the *favorite* button to the *like* button and replaced a *star* with a *heart* symbol to provide a better understanding to users. The company claims that a heart symbol is more expressive and could convey more emotion. However, many users disagreed as they tended to use favorite as a more versatile feature e.g., to save a tweet for later use or show agreement. The authors show that the *perceived affordances* are highly coupled with the name and appearance of the features and suggest that the analysis of affordances should be conducted from a user-centric perspective to obtain the precise idea of the range of possible actions people have on them. Zheng and Yu (2016) examine the uses of *Weibo*, a Chinese social media similar to Twitter, for operating and organizing Free Lunch for Children (FL4C) campaign. The authors first identified three core processes of collective actions: *construction of networks*, *framing collective action* and

*establishing legitimacy*. Then, they identified the possibilities for collective action afforded by Weibo, and discussed how they played a role in successfully driving the program by mapping each of them to one of the core practices. Again, they emphasize that the perspective of affordances-for-practice would be most relevant when using it to examine specific users within particular social, cultural and historical contexts. In the teaching-learning environment, Wang et al. (2012) conduct a cross-case analysis on three case studies to demonstrate three types of Facebook affordances: *pedagogical*, *social* and *technical*. Pedagogical affordances refer to the characteristics of Facebook that support learning activities such as sharing ideas and resources. Social and technical affordances are more general; the former supports interactions between users, e.g., between students, and between students and teachers, and the latter refers to the usability of the tool. For Facebook, although most participants found it simple and easy to use; the authors suggest that with the site being rapidly changed, some users might have difficulties using or navigating through the site.

Some prior works seem to adopt the theoretical lens of Affordances, although they do not state so specifically. Wang et al. (2016) draw on interviews with *WeChat* users to develop a model of *space-collapse*, or the emergence of public, private and parochial social spaces. Specifically, the authors examine each of the three social spaces through an examination of WeChat's technical features for user interactions. For example, WeChat offers *Look Around* to afford interactions in the public social space where individuals, in co-presence, do not know one another. A study on resilience uncovers the development of social infrastructure that supports the process of becoming resilient after crises (B. Semaan & Hemsley, 2015). They suggest social infrastructure is an assemblage of technological tools e.g., Facebook, Skype and Instant Messenger (IM) that together afford the building of resilience along four aspects: *social*

*redundancy, social diversity, developing new networks and developing trust.* The authors also note some suggestion to incorporate technical features to better support the needs related to resilience e.g., a self-identified tagging mechanism for social redundancy and diversity. Semaan et al. (2014; 2015) are particularly interested in the interactional affordances of social media for political deliberation. The studies suggest people use multiple sites, with varying affordances, to overcome the constraints of some sites. The authors derive the set of affordances necessary for user interactions in the context of political public spheres, some of which are already afforded by the existing technology. For example, the ability to aggregate information (e.g., hashtags), ability to adjust identity (some blogs allow posting anonymously) and ability to assess the impact of the content (e.g., Facebook shows number of views for videos). Interestingly, some informants reported a workaround solution to overcome the site's constraint without switching the platform, such as using dummy Facebook accounts to go anonymous.

As noted earlier, although affordances theory provides a useful foundation to assess information environment from the action possibilities perspective, the theory, however, does not offer a systematic way to examine an ability of affordances in helping users achieve the desirable results. This could be problematic especially when an object can be appropriated in many different ways. As such, I also adopt the notion of *richness* from the organizational communication literature, discussed in the following section.

### **2.2.2 Media Richness Theory**

Media Richness Theory was emerged from a study on Information Processing Theory, developed to explain the goals of communication in organizations from the information-centric

perspective (Daft & Lengel, 1986). It has generated a substantial body of literature in many different areas (e.g., organizations, friendship development and social media uses), some of which posed a challenge to the theory with contradictory evidences. The theory provides a conceptual link between managerial media choices and task performance. Daft and Lengel (1986) argued that managers choose media with suitable richness to achieve communicative acts necessary for the tasks. In other words, media choices are primarily based on the matching of *media richness* and information needs. They claimed that the communication would be more effective when task-media fit occurs i.e., the richness of a medium matches information needs.

The theory explains that managers process information to minimize *uncertainty* and resolve *equivocality* (Daft & Lengel, 1986). Uncertainty refers to the lack of information. It is the gap between information currently available and information needed to accomplish the tasks. Consensual understanding about problem interpretation already exists, filling the gap by acquiring and analyzing information will thus reduce uncertainty. Equivocal tasks are the situations which are not consensually understood. Multiple and/or conflicting interpretations exist, and acquiring more information alone may not lead to the consensual understanding.

The theory also provides a conceptual link between managerial media choices and task performance. Daft and Lengel (1986) argued that managers choose media with suitable richness to achieve communicative acts necessary for the tasks. In other words, media choices are primarily based on the matching of media richness and information needs. Specifically, lean media are preferable for uncertainty tasks (i.e., lack of information) as rich media are unnecessary and may even introduce equivocality to the communication. On the other hand, rich media are needed for equivocal/ambiguous tasks (i.e., lack of understanding) to help managers

exchange rich information, to share, and to modify subjective views until consensual understanding is attained.

To evaluate communication media, Daft and Lengel explained that media possess a set of communicative affordances that determine their capacity to carry rich information (Daft & Lengel, 1986). Such affordances contribute to medium's ability to transmit rich information in four aspects: feedback immediacy, support of multiple cues (e.g., gestures and facial expression), ability to convey natural language in addition to numeric information, and personal focus. On the basis of differences in their support of richness attributes, Daft and Lengel arrayed traditional media along a continuum describing their relative richness. Face-to-face communication is ranked highest on media richness scale. It allows rapid feedback, multiple cues to convey meanings, uses natural language and convey emotions. Face-to-face is followed by telephone, addressed written documents and unaddressed written documents (e.g., fliers and circulate letters). The table below lists the four media in an order of their richness and supports of richness attributes. This shows a simple relationship that the more a medium supports these affordances, the higher its position on the richness scale.

**Table 2.1** Media richness and traditional media. This table shows the relationship between the richness and a medium's support of the communicative affordances.

	<b>Multiple cues</b>	<b>Feedback immediacy</b>	<b>Personal focus</b>	<b>Natural languages</b>
Face-to-face	x	x	x	x
Telephone		x	x	x
Addressed written documents			x	x
Unaddressed written documents				x

The conceptual framework of media richness has been investigated in a number of ways. It was generally supported when tested on so-called traditional media i.e., face-to-face, telephone, addressed and unaddressed written documents (Daft, Lengel, & Trevino, 1987; Lengel & Daft, 1989; Russ, Daft, & Lengel, 1990). Lengel and Daft (1989) stressed the importance of the relationship between richness matching and task performance. Effective communication depends on more than using the right words to describe something or reading messages carefully; it also depends on the selection of a medium that has the capacity to engage all communication partners in mutual understanding of the message at hand.

In support of this, Daft and Lengel (1987; 1989) examined the relationship between managerial media choices and performance evaluation results. They found that managers who

were more sensitive to the matching of media richness and information needs, performed better than less sensitive managers. Russ et al. (1990) supported the notion of media selection by showing that managers tended to choose face-to-face for equivocal tasks and written documents for uncertainty (i.e., clearly understood or objective) tasks. They also supported the argument of richness matching and task performance by showing that high performing managers tended to choose the *right media*, or media with suitable richness for the tasks, regardless of their educational level, experience with organizations and introversive/extroversive personality.

Although Daft and Lengel (1986) did not incorporate so-called new media e.g., electronic mail (email) in their original study; they, later on, suggested adding email to the richness scale between telephone and written documents (Daft et al., 1987). The inclusion of new media is essentially where inconsistent findings were reported (A. R. Dennis & Valacich, 1999; M. El-Shinnawy & Markus, 1997; Fulk & Boyd, 1991; Rice, 1992; Trevino, Lengel, & Daft, 1987). These studies, however, only tested the relationship between richness perception and media choices. The common finding is while people generally perceive media richness as predicted by the theory, they do not always choose new media accordingly. Dennis and Kinney (1998) examined the selection between teleconference and text-based computer mediated communication on the decision making tasks. The results showed that people perceived teleconference richer than text-based communication as predicted by media richness theory. Suh (1999) found that the subjects tended to perceive media richness as predicted by the theory i.e., face-to-face being the most rich, followed by video, audio, and computer-mediated text. Kishi (2008) studied richness perceptions and how managers chose media. She examined both traditional media (e.g., face-to-face, meeting and telephone) and new media (e.g., teleconference,



videophone and e-conference). The results indicated that richness perceptions were consistent with the theory.

The aforementioned studies show that the subjects perceived media richness as the theory predicted. However, the selections of media were not always driven by the matching of media richness and tasks as stated by the theory. Specifically, while media choices were strongly related to the richness perceptions for traditional media, it was not the case for electronic media. El-Shinnawy & Markus (1992; 1997) argued that voice mail (vmail) was richer than email as it supported more cues e.g., tone of voice. As such, vmail should be chosen over email for equivocal tasks according to the theory. The results did not support this. They suggested, for new media, media choices were not solely driven by the media-task matching but also communication mode (i.e., textual and verbal), documentation capabilities, and user's role as sender or receiver. When the tasks – regardless of their uncertainty and equivocality levels – involve numerical information, textual communication mode is usually preferable. Documentation capabilities refer to the archiving functionality of a medium. The role of users as a sender or receiver also plays a role in media choice. The study found that receivers preferred email because its visual nature made it easier to scan quickly across and within messages.

A number of studies have found that while the affordances of a medium are fixed, richness perceptions change over time and vary by individuals (M. El-Shinnawy & Markus, 1997; Fulk, Schmitz, & Steinfield, 1990). Although, the perceptions usually converge among closely connected co-workers or cohesive groups (Ryu & Fulk, 1991). On this basis, scholars have expanded the theory by moving towards a more subjective view of richness perceptions. Fulk et al. (1990) suggested that richness perceptions were neither objective nor subjective. They are, in part, socially constructed. While they are determined to some degree by their objective

capability as explained by the theory, they are also determined to a substantial degree by social factors such as attitudes, norms and values shared by a group (M. El-Shinnawy & Markus, 1997; Markus, 1994; K. S. Suh, 1999).

People learn from their experience; the more familiar users are with a particular medium and the context, the higher richness they perceive (Carlson & Zmud, 1999; Alan R. Dennis & Kinney, 1998; M. El-Shinnawy & Markus, 1997). Walther (1992) argued that the effects of the lack of nonverbal cues of computer-mediated communication would be diminished over time as people interacted more on the medium. We can expect that people would perceive a medium as richer as the effects are diminished.

The central thesis of ranking and evaluating media by their richness is a promising idea. However, a single yardstick of media richness probably oversimplifies the complex cognitions of how people perceive new media as it fails to capture the ways that new media stretch old constraints (Carlson & Zmud, 1999; M. El-Shinnawy & Markus, 1997; Fulk & Boyd, 1991; Kinney & Watson, 1992). Although Media Richness theory had identified four attributes of the richness construct, in the end, it only focused on the broad construct, not the details (M. El-Shinnawy & Markus, 1997). For new media, some of those characteristics are more advanced and cannot be examined the same way we do with traditional media. For example, Media Richness theory predicts email as leaner than face-to-face because of its asynchronous nature. Specifically, the theory claims email supports less feedback immediacy than face-to-face, but email, in fact, can provide rapid feedback too. The inconsistency in the theory's predictive power for new media essentially led scholars to a new direction of revising attributes of richness construct.

New media come with a number of affordances which do not exist in traditional media. El-Shinnawy and Markus (1997) argued that traditional media were mostly evaluated by their capability for pushing information. However, new media come with digital asset property – for example, information can be archived and searchable. In other words, they need to be evaluated by their capability to pull information as well. While vmail is richer than email based on the four attributes, the rank could be altered if we incorporate the archival property or searchability. This explains El-Shinnawy and Markus's (1992; 1997) findings that email was chosen over vmail for equivocal tasks. Being able to see message within the context of its email thread (i.e., a series of exchanged e-mails) indeed enhances its richness, and this aspect is not captured by the original Media Richness theory. Examples of studies incorporating new attributes are Markus's (1994) study which added three attributes for examining new media: multiple addressability, external recording, and searchability, and Kishi (2008) who added two more attributes, reliability and ease of use, and removed the support of natural language attribute. These two studies show a promising area for improvement of the theory.

The wide range of affordances offered by new media introduces variety in media uses. Within the same medium, people use email, for example, in many different ways. People can make a particular email richer by attaching images or using emoticons (Brinker et al., 2015). This suggests that the richness of a medium should no longer be a distinct objective value but varies by how it is appropriated. In other words, richness should be measured on *media uses* rather than media e.g., a particular email vs. email communication.

The more recent work moves towards the direction of investigating richness within media (Sheer, 2011). The author examined the relationship between the richness of *MSN* – an instant message service operated by Microsoft – and online friendship development. She claimed that

the richness of MSN varied by how it was appropriated. She discussed that although MSN's capacities could not convey as much rich information as face-to-face communication, many of its features allowed users to communicate in thorough and multifaceted manner. She conducted a survey to examine richness perceptions of MSN's features uses, and as she notes, richness must be measured from the user perspective. The results show that the uses of webcam, MSN Spaces, animations, and icons are perceived as rich, and text message exchanges are perceived as lean. The study then examines the roles of rich and lean uses of MSN in friendship development. The results show positive relationships between rich uses and making new friends; they are useful for getting to know new people quickly but superficially. Lean appropriations are positively correlated to the deepening stage i.e., building close friendships. Simon and Peppas (2004) examines the effects of website richness on user's attitude and satisfaction. They operationalized the task of delivering information about complex products (e.g., automobile) as high equivocal tasks and simple products (e.g., audio CD) as low equivocal tasks. The hypotheses are the use of lean websites is related to positive attitude and satisfaction for simple products, and rich websites are in favor of complex products. The 2x2 experiment was set up with two versions of websites: rich and lean sites for two products: simple and complex. The rich websites present the information with both text and multimedia e.g., images, video and animations. The lean websites only present the textual information. The results support the second hypothesis of the use of rich websites for complex products. Users did not find lean websites satisfying even though product information was simple. They discussed that the advance in technology has trained users to demand richer content and presentations.

Living in a highly interactive media environment has changed our idea of what constitutes lean and rich media. Coyle and Thorson (2001) proposed that media richness is a

crucial element of creating the feeling of telepresence – a primary goal of marketing websites. Rich content (e.g., video, audio and animations) enhances the richness perception by enabling multiple senses. The experiment was conducted on different versions of websites varying in the degree of interactivity features and multimedia uses. The results show that rich websites help promote user's attitude towards the sites and a higher level of perceived telepresence as a consequence. Pollach (2008) investigates the richness of consumer opinion websites. She argues that richness perceptions varied by contexts. With the same set of affordances, the perceptions of richness depend on their appropriateness to perform a given task. For consumer opinion sites, information search is a high equivocal task, and it thus needs different set of affordances from low equivocal tasks such as review writing. She identifies and matches website's affordances to each of the four attributes of richness construct. Feedback immediacy is supported by the uses of reply, comment, and company's rebuttal to customer's review. Cue multiplicity is enhanced by an ability to view user's personal information and status e.g., credential rating. Natural language is enabled by the uses of text. Personal focus is supported by emotive icons. Du and Vieira (2012) use Media Richness theory to evaluate Cooperate Social Responsibility (CSR) campaigns on websites of oil companies. Given that CSR information is value-laden and highly complex, media richness is thus an important element for CSR communication. They measure website richness with the presences of video, image, and textual data. They found that media richness, in part, enhances the effectiveness of CSR communication.

Media Richness scholarships have extended to study social media sites. Social media are a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 technology. They allow users to create and exchange User Generated Content (UGC), which can reach people multiple times, from multiple sources, and in multiple settings

(Kaplan & Haenlein, 2010). As a result, social media have increased in popularity as an alternative communication platform.

Scholars adopt a broad view of Media Richness Theory to examine social media in various contexts. A group of studies relies on Simon and Peppas' (2004) operationalization of the uses of multimedia as rich media (e.g., Gao, 2016). This approach, however, understates the diversity in social media capabilities. In an attempt to exhaust examinations, another research area develops a classification framework based on the theory of Media Richness, and other theories, to assess the appropriateness and potentiality of social media for communication tasks. The common approach is to divide social media sites into different types by their characteristics, then classify them at the type level. For example, Ledford (2012) developed a model emerged from Media Richness and Media Control Theory – a theory that looks into the extent in which organizations can regulate the design of content and flow of information. The model was developed to select media for social marketing campaigns. Christy Ledford (2012) analyzed a number of social media types e.g., online video sharing, and microblogs. Her model uses four richness attributes of Media Richness Theory – cues, feedback, personal focus, and natural language, and two attributes of Media Control Theory – message control and delivery control. Consistent with previous work, the study notes that the perceptions or interpretations of richness might vary by audience. As such, it is imperative that the framing should be made from the audience perspective. Kaplan and Haenlein (2010) develop a two dimensional conceptual classification scheme for *social web* based on its two key elements: media-related component and social dimension. The media aspect considers the amount of information being transmitted to communicate. This element was examined using Media Richness and Social Presence Theory (Short, Williams, & Christie, 1976). Social presence refers to the capability of media to create

the sense of *being there* through the transmission of acoustic, visual and physical contact. The degree of social presence depends on the intimacy (i.e., personal vs. mediated) and immediacy (i.e., synchronous vs. asynchronous).

In an online world, social presence is gained through media richness (Lange-Faria & Elliot, 2012). Specifically, cue multiplicity attribute of media richness directly implies the intimacy attribute of social presence, and feedback immediacy is in line with social presence's immediacy. Along this dimension, they defined three levels of support: high, medium, and low. The social dimension was examined using self-presentation and self-disclosure (Goffman, 1959), and it was broken into two levels: high and low. Self-presentation explains that people have the desire to control the impressions others have about them. This is usually done through self-disclosure or the conscious or unconscious revelation of personal information e.g., thoughts, feelings, and opinions. They extended their framework to include microblogs to the classification scheme in their later work (Kaplan & Haenlein, 2011). Microblogs stand between traditional blogs and social networking sites in terms of media component and are high in social dimension. They thus placed microblogs with blogs. The modified classification scheme is presented below.

**Table 2.2** Kaplan and Haenlein's (2010, 2011) classification scheme.

		<b>Social Presence/Media Richness</b>		
		<b>Low</b>	<b>Medium</b>	<b>High</b>
<b>Self- presentation/ Self-disclosure</b>	<b>Low</b>	Collaborative projects	Content communities (e.g., YouTube)	Virtual game worlds
	<b>High</b>	Blogs and microblogs (e.g., Twitter)	Social networking (e.g., Facebook)	Virtual social worlds

This classification scheme has been adopted as a lens for choosing the platforms for communicative tasks. Ahmed (2012) argues that the communication in disaster management could benefit from both lean and rich media. He identifies three types of communication tasks based on the interactions between agency and communities: agency-to-agency (AA), agency-to-community (AC), and community-to-community (CC), which varies by the information requirements. While AA is an uncertainty task, the other two are equivocal tasks. He examines task-media matching based on Kaplan and Haenlein's (2010) classification scheme. For example, an uncertainty AA task was matched to lean social media (e.g., collaborative projects and blogs).

The classification scheme was also adopted in an online classroom study (Dao, 2015). The study identifies five characteristics of online classroom supported by the Internet: participation, openness, conversations, communities, and connectedness. The author discusses that even the leanest media were still rich enough to support all characteristics. This supports the



previous argument that for new media, and even the less rich media are still rich enough for a moderately equivocal task (Alan R Dennis & Kinney, 1998). Tsikerdekis and Zeadally (2014) adopt the classification scheme to study online deception. They note that media richness played an important role in determining the difficulty of the deception. For example, deceiving through lean media, such as texts, was found more stressful than using avatar chats but has more chances of success. Importantly, the study points out the needs to reconsider or redefine the measurement of communicative cues, as the online presence of some cues might be different from their offline forms, or do not exist. For example, Jalonon (2014) suggests social media are relatively lean, comparing to face-to-face communication, due to the lack of social cues. They, however, can modulate human collective emotion through the spread of emotionally motivated information in a way that cannot be done offline. Another view from the context of interpersonal relationship developments suggests that it is the lack of social cues that allows greater self-disclosure, which then turns the communication into unusually intimate and hyper-personal (Pollet, Roberts, & Dunbar, 2010).

The literature discussed above suggests that, when it comes to new media the *richness perceptions* are consistent with the theory. Yet *media choices* are not solely driven by the perception. That is, the richness measure alone is insufficient especially with the advancements in new media which bring about variety in media uses. Specifically, within the same medium, people use email, for example, in many different ways – for example, people can make a particular email richer by attaching images or using emoticons (Brinker et al., 2015).

As such, this study only adopts the notion of *richness*, or the ability of a medium to transmit information for solving the communicative goals and that the ability is varied by its support of communicative artifacts. Integrating with the Affordances Theory, I propose a notion

of *affordance richness* or the ability of a post to deliver information along a certain affordance dimension. More details are presented in below.

### 2.2.3 Summary

The theory of Affordances provides a useful foundation for assessing information environment from the action possibilities perspective. A relational property of affordances also suggest that an analysis of any information environment should be contextualized to the setting of the study (Fayard & Weeks, 2014; Zheng & Yu, 2016). Through its theoretical lens, this study is contextualized in relation to higher-level patterns of user behaviors and, consequently, reflects the relationships between the technological artifacts and users as afforded by environment like social media.

While the literature generally uses the theory of Affordances to investigate the affordances of a platform – for example, Wang et al. (2012) adopted the theory to study Facebook, I suggest that the theory can be adopted to examine *technologies in practice*. In other words, the unit of analysis becomes the daily practices of technology uses. This is because the technology can be used in many different ways (Brinker et al., 2015). Moreover, the theory does not offer a systematic way to examine how the affordances, when undertaken, enable users to engage in social actions. For Gibson, affordances always exist regardless of user perceptions or appropriations (Gibson, 2014a) – for example, a cellphone always affords archivable communication (via texting) and rapid communication (via calling) although an elderly person might not appropriate the archivable-communication affordance but a teenager might do. The original theory does not provide the ways in which we could use to explain the ability of this

particular use of the cellphone in helping users achieve the goal along the two affordance dimensions (archivable and rapid communication).

In searching for the ways to tackle the challenges mentioned above, I came across the theory of Media Richness. While the core concept of Media Richness is a promising idea and generally supported when testing on traditional media, it is only partially supported when testing on new media. Specifically, previous studies show that people perceive the richness of new media as predicted by the theory but do not necessarily choose media accordingly. As such, this study only adopts the notion of *media richness*, or the ability of a medium to transmit rich information which is varied by its support of communicative artifacts, and propose the notion of *affordance richness*. Similar to the concept of media richness, *affordance richness* measures the richness of affordances made possible by a medium (a post). Another way to say this is that posts with affordance richness have the ability to deliver the information necessary in affording a particular action by using some artifact. For example, when a celebrity creates a tweet (a medium) with an @mention (an artifact) to interact (an affordance) with someone in the audience, I would say that the tweet is rich in interaction affordance, or that the tweet has the ability to deliver the information necessary in affording interactions through the use of @mention artifact. In this example, I would be measuring the richness of the interaction affordance. Simultaneously, I could also be measuring the richness of other affordances (i.e., identity and visibility affordance). The concept of affordance richness could be particularly useful when an object (an environment like Twitter) can be used in many different ways. For example, Twitter always affords interaction and information-searching although I might create a tweet rich in the interaction affordance by using @reply artifact. For this particular tweet, affordance richness can be used to explain that the tweet is rich in the interaction affordance.

Next section, I present a review of three bodies of social media literature. First is a collection of studies related to the technological artifacts/functionalities of social media. Note that I will be using the two terms – *functionalities* and *artifacts* – interchangeably. Second, I present a growing body of literature related to the followership as one of the highly visible signs of celebrity status. Lastly, I present a literature concerning follower engagement – the other highly visible sign of celebrity status.

The review pays attention to Twitter and Instagram. As noted above, these two platforms are important venues for developing celebrity-fan relationships (Ward, 2016). I also suggest that they are a good point of comparison due to the differences in nature. Specifically, Twitter is a textually driven and Instagram is visually driven, but both are relatively *micro-content* service and do not employ a reciprocal relationship-based dynamic.

### **2.3 Technology Artifacts, Engagement and Followership on Social Media**

Social media sites have changed the way we communicate. They have emerged as a new key medium for information sharing by enabling people to share opinions, content, experiences, and insights through User Generated Content (UGC), which results in a continuous stream of information, opinions, and emotions (Papacharissi & de Fatima Oliveira, 2012). This trend has transformed celebrity practices towards a self-governed model, known as *microcelebrity*.

Celebrities have embraced many different platforms, each of which differs by nature, functionalities and users that inhabit it. However, little is known about the roles of such services, and particularly their communicative affordances in the practice outcome – celebrity status. The general question behind this inquiry is: *How do celebrities use social media to grow and*

*maintain celebrity status?*

Accordingly, this section presents three collections of social media studies. First, technological artifacts – as an outlook to study the uses of the platforms. Second, followership as one of the two proxies of celebrity status. Lastly, I present a review of follower engagement – the other proxy of celebrity status.

Both Twitter and Instagram do not employ a mutual relationship-based dynamic. For example, user A can follow user B, but user B is not required to follow user A. On the news feed, users see posts from a set of users they elect to follow. Twitter is a microblogging service, and Instagram is its visual counterpart. As noted earlier, they share a similarity as a relatively *micro-content* service i.e., short text or image (or short video) sharing platform. Both of them provide many different functionalities with some overlapping and others are exclusively available on one platform. While different sets of functionalities enable celebrities to appropriate the platforms in different ways, they introduce different limitations that prevent the celebrities from engaging in particular activities. In the next section, I present a review of social media functionalities, or technological artifacts, from HCI literature.

### **2.3.1 Technological Artifacts**

As discussed earlier, *affordances* are a range of action possibilities that allow users to perform certain actions by using *technological artifacts* (Kaptelinin & Nardi, 2012). The functionalities, or artifacts, of social media have been gradually developed and integrated to their architecture over time by user conventions. Some of which are common across platforms and others are exclusively available on a particular platform.

As a microblogging service, Twitter is a textually driven platform. Users construct a textual update under a 140-character limit, with an option to provide an information resource through URLs, photos, and videos. Instagram, on the other hand, is a visually driven platform. Users compose a post with a photo or short video, with an option to provide a caption text along with the post.

Both Twitter and Instagram support the use of hashtags, originally proposed to the Twitter community by Chris Messina, a software developer, as a system of channel tags for “improving contextualization, content filtering, and exploratory serendipity within Twitter” (Wikipedia contributors, 2016). Hashtags have emerged as a key feature and were integrated to the Twitter architecture and other social media sites including Instagram. They afford contextualizing the posts (Hu et al., 2014), promoting visibility (Page, 2012) by making a tweet searchable (Zappavigna, 2017), supporting trending topics (Bruns & Burgess, 2011), allowing users categorize their messages (Darling, Shiffman, Côté, & Drew, 2013), and signaling the context within which the post occurs (Golder & Huberman, 2006; Huang, Thornton, & Efthimiadis, 2010; Marwick & boyd, 2012). People use them to engage with specific topics (Bruns & Burgess, 2011; Huang et al., 2010) and for forming communities by projecting their identity as affiliated within a collective group indicated by using the same hashtag (Page, 2012).

Both platforms support actions that enable different ways that users can interact. The first of these, an @mention, can be seen as a form of *addressivity* that references others, either as the intended recipient or as a third person being talked about (Honey & Herring, 2009; Hu et al., 2014; Zappavigna, 2017). Twitter supports a conversational mechanism with @reply as a response to someone else’s tweet. On Instagram, users can comment under the posts with or without tagging other users. Both @reply and comment are intended as a discussion signal; it

may signal that a person is following along in a discussion or interested in the original content (Bruns & Burgess, 2011). Native to Twitter, retweeting is an affordance that supports human interaction. While many may dismiss retweeting as simply amplifying a message, boyd et al. (2010) suggest that retweeting is a conversational act; that users often retweet to be part of the conversation. They find that users may retweet to signal that they are listening to, acknowledging, or trying to curry favor with the person who tweeted. Retweeting also marks a tweet as worth of attention and shows an agreement with the tweet text and the user (Zappavigna, 2017). A retweet by a celebrity is an act of personal and public acknowledgement (Pennington, Hall, & Hutchinson, 2016) and creates a sense of intimacy for fans (Marwick & boyd, 2011). Seidel et al. (2016) find that people retweet as a way to associate oneself to different communities, peers, or organizations. We also know that people consider their audience when deciding whether or not to share a message into their own social network (boyd, 2008) and that they are selective about whose tweets they retweet (Kwak, Lee, Park, & Moon, 2010). Together these studies suggest that retweeting can be a complex social calculation and a kind of signaling either to the original message sender, one's audience, or both.

### **2.3.2 Followership on Social Media**

As noted earlier, one of the celebrity status proxies is followership. The number of followers one has recently become a currency of the social web (Klotz, Ross, Clark, & Martell, 2014). While existing studies primarily look at Facebook and Twitter, little is known about Instagram. Although top Instagram users are mostly traditionally famous people such as pop stars and athletes (Wikipedia contributors, 2017), some *ordinary people* also make it to the top

list, too. Examples include some Internet celebrities who have achieved status on another platform as well as some users whose fame has emerged within Instagram. Gaining followers on Instagram seems to be an ultimate goal of many users, as a large number of followers is an indicator of status (Marwick, 2015a). Marwick (2015a) conducted an interview with Instagram's founding community manager, inquiring about the methods people used to gain followers. She suggested the use of hashtags is a way to boost up the number of followers. When users embed popular hashtags, the posts appear in the explore feature of Instagram. She notes that while popular accounts tend to use only one or two hashtags, those follower-seekers tended to use a dozen of hashtags such as #followforfollow (follow for follow) to indicate their commitment to follow back.

Examining the content, Hu et al. (2014) categorized Instagram users into five groups by their majority of post content such as selfie-lover, captioned photo, and common users (posting variety of content). The study finds no relationship between number of followers and user types. That is, the following decisions on Instagram are not driven by the content of users, as measured by their categories. However, the authors did not examine the direct effects of post types on the follow decision as they classified users by majority of their post content. This could be problematic especially when considering that none of the users has only one type of messages.

A much larger body of literature examines followership on Twitter. Twitter users with many followers are often considered more powerful as their tweets diffuse much faster and wider in the network (Xu et al., 2013). Building an audience of followers can create access to a network of social ties, resources, and influence (Wang & Kraut, 2012). In recent years, the number of followers has become the most important status symbol of Twitter users. Rapid follower growth may be an early indication of a rising star, or an emerging leader, within the



network (Hutto, Yardi, & Gilbert, 2013).

While following someone on Twitter is as simple as one-click, an average number of Twitter followers in 2008 is only 85 (Huberman, Romero, & Wu, 2008), and the decision to follow someone is far from random (Kivran-Swaine, Govindan, & Naaman, 2011). The followers of any users reflect various types of relationships such as friendship, kinship, common interests, attention, or information exchange (Kivran-Swaine et al., 2011). Social media scholars have identified a number of variables which have effects on the follow decision. Lampe et al. (2007) found the number of followers correlated to user's trustworthy identity. For this study, an important signal of user's trustworthiness is the completeness of profile content e.g., whether or not URL and description are provided.

Another group of variables relate to tweet content (Kivran-Swaine & Naaman, 2011; Naaman, Boase, & Lai, 2010; Wang & Kraut, 2012). Scholars consider the content in many different ways. Examples include topical focus (Wang & Kraut, 2012), sentiment and subjectivity (Kivran-Swaine & Naaman, 2011; Quercia, Ellis, Capra, & Crowcroft, 2011), and message focus (Naaman et al., 2010).

Firstly, topical focus is the similarity of tweet topics between users. The principal of homophily asserts that similarity engenders stronger potential for interpersonal connections (McPherson, Smith-Lovin, & Cook, 2001). This suggests that the follow decisions could be driven by topic-homophily (Wang & Kraut, 2012; Weng, Lim, Jiang, & He, 2010), and Twitter users who discuss a wide range of topics may have a higher chance of gaining more followers as they appeal to a broader audience. Although the diversity in tweet content helps one gain more followers at the beginning, research found that their followers tend to be more heterogeneous

(Wang & Kraut, 2012). This is particularly useful in the context of developing online social groups. The authors identified 480 newly created accounts who self-identified as providers of political tweets, and collected their first 150 tweets. A year later, they took a snapshot of their followers and followings as the measures of accounts' success. For each user, they examined topical focus by calculating pairwise cosine similarity of vocabulary in their first 150 tweets. Specifically, it is a measure of similarity between two vectors of words i.e., two tweets, by calculating the cosine angle in a high dimensional space. High cosine similarity between two tweets indicates high text similarity, or narrow topical focus. Using negative binomial regression, they found that the initial topical focus, or similarity in the first 150 tweets, had a large impact on constructing a robust community indicated by a strong social tie amongst the members.

Sentiment matters. Research found a correlation between the expressed emotion in tweets and follower network (Kivran-Swaine & Naaman, 2011; Quercia et al., 2011). Kivran-Swaine and Naaman (2011) coded a large corpus of tweets for the presence of emotion as joy, sadness or other. The study distinguishes regular tweets from interaction tweets (@mentions and @replies) to signify two distinct types of activities which may correlate to social network properties in different ways. They constructed three models for predicting number of followers, network density, and reciprocity rate. The analyses were conducted at the user level, meaning they calculated the proportions of posts expressing joy, sadness, and other to the total posts in each category for each user. Using stepwise regression analysis, they found the expression of emotion in interaction tweets was significant in predicting number of followers, but it negatively affected network density. In other words, the expression of emotion is associated with sparser network.

Interestingly, one would expect people to share emotional content with their close ties but the results suggest differently. Quercia et al. (2011) examined whether different types of users

used language differently in their tweets. They identified five types of users: *popular* (measured by number of followers), *influential* (measured by being mentioned and retweets), *listeners* (follow many users), *stars* (being followed by many users), and *highly-read* (being listed by many users). Using a standard dictionary, they categorized tweets into 72 types e.g., positive and negative emotion, work-related, and cognitive processes. Using correlation analysis, they showed that popular users predominantly expressed positive emotions in their tweets, concerned with one-on-one as indicated by the use of second person pronouns, and tweeted about the *here and now* as indicated by the use of present verbs.

Another aspect of tweet content is the nature or focus of messages. (Naaman et al. (2010) identified two broad categories of Twitter users as *Meformers* and *Informers*. The first category, *Meformer*, refers to the users whose majority of tweets focus on the self and are more personal in nature e.g., tweet about oneself or one's own thoughts. The other category, *Informer*, refers to those whose majority of tweets are more about the dissemination of informational content. They found that about 80% of Twitter users were Meformers, but those in the smaller group of Informers had far more followers. They also found that Informers used @mentions more frequently. The research suggested that Informers had more interesting content and therefore attracted more followers. An alternative explanation is that an increase in followers encourages user to post additional (informative) content. However, the authors did not examine the direct effects of tweet nature (either Informer or Meformer) on the follow decision as they classified users by majority of their tweets. Similar to Hu et al.'s (2014) study on Instagram, this could be problematic especially when considering that none of the users has only Informer nor Meformer messages.

A more recent study offers a longitudinal analysis of the changes in followers over time (Hutto et al., 2013). The authors developed a large corpus of tweets from 507 active users over the period of 15 months, and took snapshots of users' followers and following every three months. The study constructed a negative binomial model to test the relationship between the predictors and follower growth. They highlighted the importance of message content variables on follower gain. The significant variables are sentiment, informational-focus i.e., Informer message (Naaman et al., 2010), number of retweets, hashtags, and linguistic sophistication. The study offers a great starting point for longitudinal studies, however, the authors examined all predictors as an aggregated value over a three-month period. I suggest that there could be some effects at the tweet level, as evidenced by Quercia et al. (2011) that the expressed emotion was significant for interaction tweets (indicated by @mentions) but not regular tweets. Additionally, aggregating all tweets within a period of three-month was probably too coarse to capture the real effects of tweet content on the follow decision.

An alternative approach to address the question of how to gain more followers is knowing how to maintain them. A substantial body of studies have investigated the dissolution of network ties, which occur when Twitter users decide to unfollow others (Kivran-Swaine et al., 2011; Kwak et al., 2011; Xu et al., 2013). Such studies examine two groups of factors: relational and informational factors. Relational factors are related to network structure based on users' relationships – for example, reciprocity (mutual relationships), relationship ages, social status (measured by number of followers), and common friends (relationship overlapping). Informational factors are tweet related factors such as tweet topics and content.

A large scale study of Korean Twitter users shows that both relational and informational factors are crucial for the unfollow decisions (Kwak et al., 2011). The authors collected daily

snapshots of follow relationships and tweets from 1.2 million users over the course of 51 days. The quantitative analysis indicates that both relational and informational factors were crucial in unfollow decisions. For the relational factors, the results show that reciprocal relationships are less likely to be broken, unfollow occurs more frequently with the newer established relationships and less frequently when users have more common friends. For the informational factors, the authors asserted that unfollow decisions were partly driven by tweet's informativeness. That is, users unfollow others when they no longer find their tweets interesting. They measured tweet's informativeness through retweet and favorite counts. The finding is the likelihood of unfollow is decreased when tweets are more informative, as indicated by getting retweeted or favorited. Supplementing the findings with the interview, the authors found that the most frequent reason for unfollowing was information overload, or when a user tweets too much regardless of content. Another reason was related to tweet content; people are likely to unfollow when users tweeted about the mundane details of daily life. This is similar to Naaman's (2012) Meformer category. Along the same line of the topic-homophily concept (Wang & Kraut, 2012), respondents unfollowed when the topics were not interesting to them regardless of the quality of tweets.

Another study on relational factors (Kivran-Swaine et al., 2011), collected two snapshots of follow relationships to identify what factors were crucial for unfollow decisions. Similar to Kwak et al.'s (2011) work, reciprocal relationships and the number of common friends negatively affected the unfollow decisions. The work also examined social status of users and showed that users were less likely to unfollow users who had more followers than themselves. They called this *prestige ratio*. Xu et al. (2013) note that user behaviors might differ from group to group. For example, the reason for unfollowing a friend could be different from unfollowing a

celebrity. The study focused on ordinary users or Twitter users with 1,000-2,000 followers. They took four snapshots of their followers to test the effect of relational and informational factors. The study examined three relational factors (mutual friends, number of followers and number of common friends) and five informational factors (topic-homophily, uses of @replies, @mentions, retweets and favorites). Using logistic regression with the longitudinal data, they showed that only relational factors had significant impacts on the unfollow decisions. Specifically, mutual following ties, the number of followers, and the number of common friends all have negative impact on the unfollow decisions. On the other hand, informational factors have no significant impact. This is, however, contrary to the literature which suggests that informational factors have effects on the follow/unfollow decisions. An explanation might be that tweet content was only important to relationships like celebrity-fan groups, as they are formed based on the common interest.

### **2.3.3 Follower Engagement on Social Media**

The other proxy of microcelebrity status is the degree in which ones engage their audience. While gaining followers is an ultimate goal of attention-seekers like most of the microcelebrity practitioners, engaging followers is also important as a means to maintain audience (Kwak et al., 2011) and a form of social feedback (Bakhshi et al., 2014). Although engagement could take different forms on different platforms, they are all a mechanism for followers to communicate with the poster, and vice versa, around the content. On Instagram, users engage and interact with the poster by commenting and liking posts. Likes on Instagram are regarded as a social signal of “Instagram worthy” (Abidin, 2014, p. 123), but are dispersed

i.e., majority of the posts get only few likes (Araújo et al., 2014; Bakhshi et al., 2014). Twitter also has the like feature (favorite), even though people tend to use retweets and @replies more often. With multiple options of showing engagement or responding to the posts, there is no standard engagement measure (Vadivu & Neelamalar, 2015). To the best of my knowledge, none but Facebook has revealed its official engagement formulas (Facebook, 2017):

**Equation 2.1**      
$$Facebook\ Engagement = \frac{\# Likes + \# Comments + \# Share}{Post\ Reach}$$

The formula can be applied to Twitter and Instagram by substituting *Post Reach* with number of followers and altering numerators as appropriate. For example, they should be *likes* and *comments* for Instagram, and *likes*, *replies* and *retweets* for Twitter. This section presents previous work looking at different forms of user engagement: likes, replies (or comments), and retweets.

Research on Usenet newsgroup, a discussion forum that allows users to post and comment, found that both content and the posters were all affected the probability of getting replies. For example, new users are less likely to get replies than the established members (Arguello et al., 2006), and politeness has different effects on number of replies in different groups (Burke & Kraut, 2008). Closer to Instagram is Pinterest, a photo-sharing platform that allows users to *pin* photos they found online and categorize into collections, where other users can *re-pin* (share), like, and comment on photos. Gilbert et al. (2013) investigated both user and content factors that had effects on getting re-pins from other users. Using negative binomial regression, they showed that female users tended to get more re-pins, and that users tended to get less re-pins as they created more posts. On Instagram, Bakhshi et al. (2014) collected a million of posts to examine the effects of the presence of face on getting likes and comments. Using a face

detection module and negative binomial regression, they found that the posts with faces received 38% more likes.

Audience engagement is time-dependent. An interview with Instagram users in Singapore indicated that the best times to get likes were from 8-10am and 7-9pm weekdays (Abidin, 2014). Similar to follower-seekers, hashtags are an important mechanism to boost up the number of likes. When users embed a popular hashtag (either global e.g., #ootd or outfit of the day, or personal hashtags of some popular users), the post will appear in the explore feature of Instagram. Some users strategically combine such popular hashtags with a personal hashtag in order to gain visibility for their post and the personal tag at the same time. The global tags will get the posts to appear in the search feature. On the search page, other users will be tempted by a personal tag, that will lead them to a personal stream that achieves all the posts with this particular tag (Abidin, 2014). Similar to other social network platforms, the number of likes follows a power-law distribution. Within a collection of 1.2 million Instagram posts, about half received no like at all (Araújo et al., 2014). This study also noted an importance of using hashtags related to current events or celebration dates as a way to show user's reaction to the events, which leads to more likes. However, using too many hashtags tends to result in less likes. Importantly, the study highlights the *rich-get-richer* phenomenon (Barabasi, 2003). That is, highly followed users tend to get more likes that could turn posts to even more popular.

On Twitter, retweets are generally regarded as a typical, but cheap form of engagement (B. Suh, Hong, Pirolli, & Chi, 2010). Retweetability is a relatively large body of literature, comparing to Instagram's likes/comments. In the most general sense, retweeting is the act of diffusing a piece of information originally developed by others. It is also a form of participating in a conversational ecology and creating a sense of community. As such, celebrities are both



retweeting others and looking forward to getting retweets (boyd et al., 2010). However, getting retweets is not easy; Zaman et al. (2010) suggest that retweeting happens when a user feels a tweet is important enough to share with his/her network. The common theme of the prior works about retweetability considers all followers of Twitter users as retweeting candidates, whose decisions are influenced by a number of factors such as the profile of a tweet's creator (Uysal & Croft, 2011) and tweet content (B. Suh et al., 2010). Collectively, scholars have identified and grouped the factors into three categories: user-based, tweet-based, and content-based factors (Hong, Dan, & Davison, 2011; B. Suh et al., 2010; Uysal & Croft, 2011; Zaman et al., 2010).

The first category is the information about Twitter users and how active they are on Twitter. Uysal and Croft (2011) suggest account age, the presence of profile's description, and the numbers of followers, friends, tweets, and favorites are all related to retweetability. The latter part is advanced in Lee et al.'s (2015) work which examines the role of self-disclosed occupation information on the influential level of Twitter users. They find that users with undisclosed occupations have more chances of producing influential political tweets with high retweetability rate regardless of their numbers of followers. Even though they could not explain the reasons behind the findings, they did find that a significant number of such users closed their Twitter accounts or hid their tweets right after the election. Retweet users also play a role in driving more retweets to the original tweets (Hemsley, 2016). Tweets are more likely to get more retweets when they are retweeted by users with the high number of followers.

Considering the uses of Twitter affordances, such as @mentions and URL, Suh et al. (2010) and Uysal and Croft (2011) show that tweets with hashtags, URLs, and @mentions are significantly more likely to be retweeted. (Uysal & Croft, 2011) suggest the uses of question marks, exclamation marks, quotation marks, and first person pronoun also affect the

retweetability. Another interesting feature is whether or not a tweet has been retweeted before as investigated by Hong et al. (2011). They construct a binary classifier to predict if a tweet will get retweeted. One of the classifier's attributes is a Boolean variable indicating whether a tweet has been retweeted before. They find that being retweeted once increases probability of tweets to get more retweets.

Content of tweets (e.g., novelty and emotions) also matters. The third group of attributes is content-based, which expresses information contained in a tweet. Petrovic et al. (2011) manually label tweets as having novel content or not and then train a Machine Learning algorithm to categorize the larger set of tweets. They found that tweets being rated as having novel content were significantly more likely to be retweeted. Using a different approach, Uysal and Croft (2011) define novelty as the distance between a tweet and other tweets in a user's timeline. That is, the more different a tweet is from others in its network, the more novel it is. Novelty seems to be less important in the global network. Yang et al. (2010) performed a content analysis and suggest that tweets about *hot topics* are more likely to get retweets. For them, the hot topics are those being frequently mentioned in the tweets corpus. Emotions expressed on tweets also play a significant role. Stieglitz and Dang-Xuan (2012) construct regression models to examine the relationship between the number of retweets and emotions while controlling for the number of followers, account age, and hashtag inclusion. They quantify emotions by counting positive and negative emotion words in tweets. The regression model suggests that tweets with negative sentiments are more likely to induce more retweets. They also found that the retweetability was higher when tweets contained words that reflected affective dimensions such as by associating with certain political parties or politicians.

### **2.3.4 Summary**

This section presents a collection of social media studies in three aspects: technological artifacts, followership, and follower engagement. The literature shows that followership and engagement are inextricable phenomena where one could play a significant role on the other (Hemsley, 2016). A tweet by highly followed users is more likely to get more retweets than tweets from others. When a tweet get retweeted, it is brought to a new audience who can potentially become a new follower of the author. Additionally, when a tweet was retweeted by highly followed users, it is more likely to get more retweets than when retweeted by users with small number of followers.

## **2.4 Literature Summary**

This chapter has presented collections of previous work in three areas as the foundation for the theoretical direction of this study. The first section presented a wide range of studies discussing celebrity culture in mainstream media, then moved toward celebrity as a practice in the age of social media using the theory of Microcelebrity. Microcelebrity is a set of self-presentation techniques engaged by both traditionally famous and ordinary people to amp-up their popularity using multiple social media platforms. In addition to the need to maintain a consistent persona, the core microcelebrity practices are identity construction, interaction with fans, and promoting visibility.

The second section presented the Affordances Theory to ground the design of a richness framework, and the theory of Media Richness, whose notion was borrowed and modified to

create a measurement of *affordance richness*. The richness scoring framework is designed for examining the ability of social media uses to afford delivering rich information and solve communication goals.

Since the research sites of this study are two social media platforms – Twitter and Instagram – I presented a collection of social media studies in three areas. The first area was a review of technological artifacts, which reflected the mediated action perspective that was adopted to study the uses of social media. The other areas, followership and user engagement, were presented to complement microcelebrity studies. Specifically, the number of followers, likes, comments (replies), and retweets are visible signs of celebrity status of Twitter and Instagram users.

## **2.5 Research Questions**

Building on the discussed literature, this study examines the richness of social media uses in the context of microcelebrity practices. Recall the overarching question regarding the uses of social media for growing and maintaining audience. Specifically, I am interested in exploring the ways in which celebrities a) develop and maintain their online identity, b) interact with fans, and c) grow their popularity beyond the existing fan base, to expand and maintain audience by examining their social media activities. In the following paragraphs, I outline specific research questions emerging from the literature, across the spectrum of fame – mainstream famous and Internet famous practitioners.

Previous work suggests that not all the celebrities would use the same mix core practices of microcelebrities to the same degree (Marwick, 2015a; Rahmawan, 2013). Given the trend of

using multiple social media sites (Greenwood et al., 2016), it could be the case that they engage in different activities, on different platforms. Although this could be challenging as they need to maintain a consistent persona. However, using multiple platforms would give them opportunities to expand their audience and overcome limitations of a particular platform. For example, Instagram is known to be limited in interactivity and Twitter is quite short on content with the 140-characters limit. In order to delve deeper into the practices in the broad media landscape, this raises the first question:

***RQ1:** Along the core practices, how do celebrities engage in different activities on different social media platforms?*

Literature shows similar and different ways mainstream and Internet celebrities engage in the practices on different platforms. For example, both of them rarely interact with fans on Instagram (Marwick, 2015a; Ward, 2016), but some studies have documented the interaction work of mainstream celebrities on Twitter (Huba, 2013; Pegoraro, 2010), but not Internet celebrities (Rahmawan, 2013). While the attention in the literature is placed on the uses of hashtags by Internet celebrities on Instagram (Abidin, 2015; Marwick, 2015a), little is known about whether or not mainstream celebrities also make use of this mechanism. This raises another question:

***RQ2:** How are the practices similar and/or different amongst mainstream and Internet celebrities engaging in microcelebrity?*

Thinking of microcelebrity as a performance which “is molded and modified to fit into the understanding and expectations of the society” (Goffman, 1959, p. 35). In this view, audience members or fans play a role in co-constructing the performance and media environment within

which celebrities operate (Papacharissi & de Fatima Oliveira, 2012; Senft, 2008; Thrall et al., 2008; Usher, 2015). As such, it is important to understand how the audiences respond to different microcelebrity strategies. Importantly, research shows that certain behaviors result in different outcomes when engaged by different actors (Araújo et al., 2014; Xu et al., 2013). I expect that an audience would react to traditional and Internet celebrities differently even though they employ the same strategies. Another question is:

***RQ3:** How do the audiences respond to different types of strategies when controlled for celebrity types?*

With the literature suggesting the public expect their celebrities be consistent online (Marshall, 2006; Turner, 2013), ones might expect that they might be less engaged and/or even unfollow if a celebrity were to be inconsistent in the ways they use social media. Another question is:

***RQ4:** How do the audiences respond to the changes in microcelebrity strategies?*

On the basis of audience members playing an important role in co-constructing the performance and media environment within which celebrities operate (Papacharissi & de Fatima Oliveira, 2012; Senft, 2008; Thrall et al., 2008; Usher, 2015), understanding expectations and behaviors of fans is important as a means to enhance the practice outcome, and sustain promotional activity (Usher, 2015). This brings the last question:

***RQ5:** Why do the audience respond to celebrities they ways they do?*

In the next chapter, I present the methodological design of the study which comprises of two sequential phases: quantitative and qualitative analyses. The quantitative component of the

work primarily relies on the richness framework. I provide more detail about how the practices of microcelebrity are conceptualized, developed, and combined to form the framework. I also present the potential designs of richness score analyses for answering the questions noted above including a discussion on data collection, analytical methods, and expected outcomes. The qualitative component aims at answering the last research question by providing causal explanations and assisting the interpretation of the quantitative results.

## **CHAPTER 3**

### **METHODOLOGY OVERVIEW**

This chapter presents an overview of the methodological model for studying microcelebrity practices on multiple social media platforms and explains the methods I used. I adopted an explanatory sequential mixed-methods design, or the uses of qualitative results for assisting the explanation and interpretation of quantitative findings (Creswell, 2013). With the primary theme of the co-construction of microcelebrity performance, both phases of the study examine the practices from the perspectives of celebrities and fans to ensure the analyses embrace both sides of the co-construction (i.e., celebrities and fans). The presentations of each phase of the study will be as follow.

The first phase of the study was a collection of quantitative analyses and consisted of two sequential parts: framework development and richness score analysis, details are presented in Chapter 4. The results from the quantitative phase, presented in Chapter 5, were organized into three main themes within the perspective that microcelebrity performance was co-constructed by celebrities and their fans. The results were then used to inform the design of the follow-up qualitative study, presented in Chapter 6. The data for the qualitative study were collected from semi-structured interviews with audience members. This phase of the study was designed to confirm, clarify and provide causal explanations about findings about audience responses to celebrities and how they supported the claim that microcelebrity performance was co-constructed by celebrities and their fans. Results are presented in Chapter 7.



This chapter first reviews methodologies adopted by previous studies. Then, I present the overview of my methodological model.

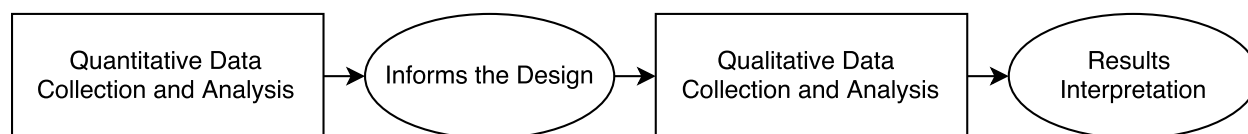
### **3.1 Methodology Review**

Microcelebrity studies have adopted a wide range of research approaches. Qualitative case studies were drawn from digital discourse analysis of text and visual content of a few users (Abidin, 2015; Bennett, 2014; Huba, 2013; Marwick, 2015a; Ward, 2016). Some studies supplemented discourse analysis with interviews (Abidin, 2014; Marwick, 2013; Mavroudis & Milne, 2016; Senft, 2008). Although this approach could provide deep insights from the practitioner's point of view, approaching them is challenging (Mavroudis & Milne, 2016) and usually results in small scale studies that limit the generalizability of the study. Mavroudis and Milne (2016) noted a challenge in gaining access to the subjects due to their closed group nature, and they rarely responded to scholars.

A relatively small number of studies adopted a laboratory experiment approach. In one notable exception, Jung et al. (2017) conducted an experiment to collect users' responses to different Instagramming strategies of the politicians. Another group of studies adopted qualitative content analysis. They analyzed either textual, visual or both forms of content to develop a codebook, then classified each post to one or more categories (Kassing & Sanderson, 2010), and drew conclusion from statistical analyses (Frederick et al., 2014; Golbeck et al., 2010; Hemphill et al., 2013). Such studies are large scale analyses with more generalizable results.

While each approach has its own strengths and weaknesses, the mixed-methods may provide a more comprehensive look and offer a more complete picture of the results through the

complementary strengths and non-overlapping weaknesses (Creswell, 2013). For my research, I choose an explanatory sequential mixed-methods design (Creswell, 2013) that starts with quantitative data collection and analyses, followed by a qualitative study to explain the findings. Specifically, the results from the first component informed the design of the follow-up qualitative study. The overviews of each phase are presented in the following sections.



**Figure 3.1** Research Paradigm (adapted from (Creswell, 2013)).

### 3.2 Quantitative Studies

The first phase of the study is a collection of quantitative analyses. I began with the design and development of a richness framework, which was then used as a tool for quantifying social media activities as measurable and comparable richness constructs. Once the framework was established I conducted a series of richness analyses, using data from both mainstream and Internet celebrities, to explore the relationships between richness measures and the outcomes of the microcelebrity practices.

This phase aims at answering the first four research questions such as the similarities and differences between the practices on Twitter vs. Instagram (RQ1) and how the audience responds to different strategies of microcelebrity (RQ3). The analytical methods include various statistical approaches – for example, tests of equal means and regression analysis. The analyses provided information about the affordance richness of a celebrity’s social media uses and its relations to

the outcome of these practices – *celebrity status*. Recall that affordance richness is a measure of the ability of affordances to help celebrities achieve the goal. The results were then used to inform the design of the follow-up qualitative study. More details are presented in Chapters 4 and 5.

### 3.3 Qualitative Study

The second phase of the study relied on qualitative methods to further uncover the relationship between affordance richness and audience responses from the audience perspective. This component of the study aims to address RQ5: *Why do the audiences respond to microcelebrities the ways they do?*

Audience's expectations and behaviors are essential for co-constructing celebrity performance and the media environment within which celebrities operate (Goffman, 1959; Papacharissi & de Fatima Oliveira, 2012; Senft, 2008; Thrall et al., 2008; Usher, 2015). In this dissertation, I operationalized the *audience* as social media users who followed celebrities and recently interacted with them on Twitter and Instagram. The data were collected using semi-structured interviews with the target group through a random sampling strategy (Robinson, 2014). The interview instrument was designed based on the results from the quantitative analyses. This study is expected to be a supplementary dataset and analysis for the results from the framework, and thus strengthen the interpretations. Details are presented in Chapters 6 and 7.

## CHAPTER 4

### QUANTITATIVE METHODS

As noted earlier, the first phase of this dissertation is a collection of quantitative studies and consists of two sequential parts: framework development and richness score analysis. The design and development of the framework was based on the theoretical lens of Affordances (Gibson, 2014a), responses from crowdsourcing annotations and machine learning models. For the richness score analysis, I used a collection of Twitter and Instagram data from both mainstream and Internet celebrities. The framework was used to quantify social media activities using affordance richness scores. The scores were analyzed with statistical approaches, such as equality of means tests and regression analysis in order to draw conclusions from inferential statistics. The results from the quantitative phase then provided information about the affordance richness of celebrity's social media uses, and its relations to the practice's outcome – *celebrity status*.

This chapter first presents the design of the richness framework which consists of two components: the structural design and the richness component. Then, I present the analytical methods for analyzing the richness scores generated by the established framework.

#### **4.1 Richness Framework**

This phase of the study was based on observational social media data. Observational data mean they are observed and collected from a sample of population who are not under the control

of researchers (Rosenbaum, 2002). They are particularly useful for providing information about real world phenomena by observing the general population. Social media data are rich sources for observing real world phenomena across different areas of social science research such as political communication (Golbeck et al., 2010; Hemphill et al., 2013), marketing (Abidin, 2015), and online learning (Dao, 2015) to name a few. The two research sites for this study are Twitter and Instagram.

Recall that the general question behind this inquiry is: *How do microcelebrities use social media for growing and maintaining celebrity status?* A challenge in answering the question is the lack of a systematic way to examine social media activities that allows an analysis to look beyond any specific platform. In the context of microcelebrity, an examination of the practices should preserve different dimensions of practice along which people might engage i.e., identity construction, interacting with fans, and promoting visibility. Literature shows that certain social media behaviors could have different effects when conducted by different actors (Xu et al., 2013). As such, the designs of any assessment frameworks should be contextualized to the setting of the study.

To systematically examine social media affordances in the context of microcelebrity, I designed and developed a richness framework that serves as a tool for quantifying social media activities to measurable richness constructs. Such constructs can be analyzed in different ways to answer a wide range of questions regarding the uses of social media in the context of microcelebrity practices. Specifically, I suggest that *information environments* of celebrities are comprised of strategic combinations of their core practices (i.e., identity, interaction and visibility), and that we can learn about their practices by studying the information environments within which the celebrities operate in using the richness framework.

In the following sub-sections, I present the structural design of the framework. I also explain how I adopted the concepts of Affordances Theory and describe the ways in which I quantified microcelebrity practices as richness scores. With the structure in place, I describe how I established the richness scoring component and formed the richness framework. Lastly, I present celebrity status measures, which include audience growth and audience engagement.

#### *4.1.1. Structural Design*

The theoretical foundation of the framework primarily draws on the conceptual lens of Affordances Theory (Gibson, 2014a). This theory provides a foundation for assessing an information environment from the affordances, or action possibilities, perspective. In the context of information systems, affordances are re-defined as “the possibilities for goal-oriented action afforded by technical objects to a specified user group by technical objects” (Markus & Silver, 2008, p. 624). That is, affordances are the *relational action possibilities* that users can perform (i.e., common usage patterns) by using the technical objects (i.e., technical functionalities/artifacts e.g., @mention and hashtags). A relational property mean that of affordances can differ by users or user groups but also suggest that an analysis of any information environment should be contextualized to the setting of the study (Fayard & Weeks, 2014; Zheng & Yu, 2016) because the affordances are “inextricably bound up with specific, historically situated modes of engagement and ways of life” (Bloomfield, Latham, & Vurdubakis, 2010, p. 415).

Through a conceptual lens of Affordances Theory, the framework examines the characteristics of social media based on the common usage patterns of users. In the context of

microcelebrity, I suggest social media sites provide three affordances mapped to the core microcelebrity practices. Previous studies have collectively suggested that the core practices of accruing celebrity status are the construction of *identity*, *interactions* with fans and promoting *visibility* to expand a fan base (Marwick, 2013; Page, 2012). The first dimension, *identity affordance*, is an ability of social media to afford users to position the self in relation to others (Page, 2012) by sharing information which reflects one's identity, or what they want others to have impression about them (Khamis et al., 2017; Marwick, 2015b). The second dimension, *interaction affordance*, affords celebrities the ability to interact with their fans and engage in parasocial relationships in the public social space. Interaction affordance allows celebrities to develop and maintain their audience by responding, or reaching out to fans as a means to create a sense of conversation (Raun, 2018). The last dimension, *visibility affordance*, affords celebrities the ability to promote and compete for public attention. Visibility affordance enables microcelebrity persons to be found by the public beyond the existing fan base in order to expand audiences and become more popular.

In this way, I can examine the abilities of social media without restricting an analysis to any specific platform, while preserving different dimensions of practices along which the practitioners might engage. Each social media platform provides a different set of technology artifacts, or functionalities, that contributes to affordances in different ways (Fayard & Weeks, 2014; Kaptelinin & Nardi, 2012; Zheng & Yu, 2016). Within each platform, users can construct a message e.g., tweet or Instagram post, in different ways using different combinations of technology artifacts. Thus, each post varies in *affordance richness* or the ability to serve the information needs of celebrities along the three dimensions of affordances (i.e., identity, interaction and visibility).

Similar to the concept of media richness, *affordance richness* measures the richness of affordances made possible by a medium (a post). Specifically, Daft and Lengel (1986) suggest an ability of a medium varies by the communicative artifacts it possesses. For example, face-to-face communication is the richest while a memo is the leanest media. The literature also suggests that an objective measure of media richness at the media level oversimplifies how people perceive new media, as the wide range of functionalities has introduced variety in media uses (Brinker et al., 2015; Coyle & Thorson, 2001; Du & Vieira, 2012; Simon & Peppas, 2004). That is, within a communication medium, richness varies by how it is appropriated. In this work, the notion of *media richness* is altered to measure the richness along the affordance dimensions or *affordance richness*. Another way to say this is that posts with affordance richness have the ability to deliver the information necessary in affording a particular action by using some artifacts of social media.

To measure the richness, the framework organizes the technology artifacts of social media into groupings by their relevance to an affordance dimension. The framework measures the affordance richness of tweets, or Instagram posts, based on the way they are constructed – the affordance richness reflects the uses of technology artifacts within the associated affordance-artifact grouping.

As affordances should be derived from the user perspective (McVeigh-Schultz & Baym, 2015), the artifact-affordance mappings were based on HCI literature concerning mediated action possibilities enabled by social media (boyd et al., 2010; Honey & Herring, 2009; Hu et al., 2014; Kwak et al., 2011), and the contribution of each artifact to the affordance dimension was derived based on the wisdom of the crowd. Learning from the crowd is particularly important when we consider that affordances are user perceptions about action possibilities enabled by technology artifacts (Kaptelinin & Nardi, 2012). For example, an @-sign is a technology artifact that affords



*addressivity*. Twitter users perceive this artifact as a mechanism for addressing other users, known as @mentions (Honey & Herring, 2009; Hu et al., 2014).

With the structure of the framework in place, in the next section I present how the richness scoring component was developed.

#### *4.1.2. Scoring Development*

The richness scoring framework has been developed through successively more sophisticated versions over time and results in three published articles. The following subsections present different versions of the framework, their limitations, and how they were refined to overcome such limitations.

The first work (Tanupabrungsun, Hemsley, Semaan, & Stromer-Galley, 2016) developed a Tweet Quality Assessment Framework (TQAF) to examine differences in the tweeting behavior of politicians while running for office vs. after holding office for six months. With TQAF, I measured three dimensions of tweet quality: *contextual*, *interaction*, *information*, and a combination of the three dimensions for an *overall* quality score. The quality score in each dimension is an un-weighted Euclidean distance of the uses of relevant artifacts. For example, the interaction quality score is a combination of the presence of RTs, @mentions, and @replies. The framework does have a few important limitations. Specifically, it lacks theoretical underpinning, and the weighting within and across dimensions is equivalent. In the TQAF, affordances within each richness dimension were weighted equally. For example, when calculating the score of the interaction quality dimension, RTs, @mentions, and @replies were all weighed equally. Early work in psychology suggests that body language, tone of voice, and

spoken words all have significantly different weightings in communicating meaning (Mehrabian & Ferris, 1967). This suggests that the three quality dimensions should be weighed differently when calculating an overall quality score, and that within each dimension, artifacts should be weighed differently. This work serves as a starting point for studying an information ecology of social practices through a perspective of technology affordances.

The second version of the framework (Tanupabrungsun, Hemsley, & Semaan, 2018) addressed the theoretical limitation by grounding the framework with a conceptual lens of Media Richness Theory (Daft & Lengel, 1986) and Affordances Theory (Gibson, 2014a). Drawing on these theories, my framework claims that actors have different needs in solving *uncertainty* (i.e., lack of information) and *equivocality* (i.e., lack of mutual understanding) and the differences are reflected by their uses of different *richness dimensions*. I used the framework to examine a corpus of Occupy Wall Street tweets and emphasized on those activities of the *core actors*, who were largely instrumental in moving the movement forward. These actors needed to engage in various information processing activities to solve the problems of uncertainty and equivocality as a means to achieve the goals of the movement. The framework categorizes the tweets into sub-groups based on their nature as reflected by the uses of technological artifacts. Although the framework was strengthened in terms of the theoretical foundation, it still has not addressed the methodological limitation about the calculation of richness scores.

In the third version of the framework (Tanupabrungsun & Hemsley, 2018), I refined the framework to use a more sophisticated richness score calculation by weighting each of the individual artifacts (e.g., @mentions, URL) differently. For this version, the richness score is a linear combination of weights and the uses of communicative artifacts illustrated in the formula

below. Each of the artifacts within a group is represented by  $v_i$ , and their corresponding weights are represented by  $b_i$ .

**Equation 4.1** 
$$richness = b_1 * v_1 + b_2 * v_2 + \dots + b_n * v_n$$

To weigh the contribution of each artifact to its richness measure, I used a combination of annotations by crowdsourcing and classification modeling using logistic regression, where the coefficients of the regression become the weights. Specifically, I utilized Amazon Mechanical Turk (AMT), the crowdsourcing service operated by Amazon, to develop a training dataset and developed a classification model to automatically annotate a bigger set of data. This work touched upon a microcelebrity literature and used the framework to examine a corpus of tweets from mainstream celebrities in different domains (e.g., pop stars and sports stars). This work offers a methodological direction for obtaining the richness annotations. However, the conceptualization of the richness dimensions as informational, interactional and contextual are too generic and the connection to microcelebrity literature needs to be strengthened.

In summary, these previous studies (Tanupabrungsun & Hemsley, 2018; Tanupabrungsun et al., 2018, 2016) have collectively informed the design of the framework used in this dissertation. Theoretically, the framework is designed through a perspective of technological affordances to examine the characteristics of an information system (e.g., social media) based on the common usage patterns of users and how they help users achieve the communicative goals. The methodology for obtaining the richness annotations can be replicated through the uses of crowdsourcing labelling tasks and classification modeling. However, the richness dimensions should be refined to better connect to the theory of Microcelebrity.

In this study, I suggest social media sites provide three affordances mapped to the core microcelebrity practices. The first dimension, *identity affordance*, is a more definite form of informational richness which affords celebrities to position the self in relation to others (Page, 2012) by sharing information which reflects one's identity, or what they want others to have impression about them (Khamis et al., 2017; Marwick, 2015b). The second dimension, *interaction affordance*, is similar to interactional richness which affords celebrities the ability to interact with their fans and engage in parasocial relationships in the public social space. The last dimension, *visibility affordance*, is a definite form of contextual richness, which affords celebrities the ability to be part of different communities as a means to gain public exposure, promote themselves and compete for public attention.

In the following sections, I present the datasets developed and employed in this study, how I obtained the annotations using the methodology explained earlier, and the modeling process to automatically generate the richness scores of the unlabeled datasets.

### *A. Datasets*

The collections of tweets and Instagram data were developed using tools that collected data from users' timelines. Each dataset is a collection of posts from mainstream and Internet celebrities. The list of users was constructed with two approaches. First, I relied on previous studies that revealed the names of celebrities (Abidin, 2014, 2015; Marwick, 2015a; Mavroudis & Milne, 2016). Second, I gathered several online lists compiling a collection of *trending users* on social media. The list of mainstream famous users contains the top pop stars, athletes and scientists who have achieved offline status, and made use of social media, based on lists curated

by The Guardian, Forbes, and Science. This process gave a total of 90 names. For Internet celebrities, the list was consolidated from the lists by Forbes<sup>2</sup>, Elle<sup>3</sup>, Marie Claire<sup>4</sup>, Pop Crunch<sup>5</sup> and Greatist<sup>6</sup>. This process gave a total of 105 users and comprised users from different domains such as entrepreneur, fashion, and fitness.

Then, I went over each user and searched for his/her Instagram and Twitter accounts. A candidate was added to the final list if he/she had public accounts on both platforms, each account was still active, and posts were in English. This process produced the final list of 33 traditional celebrities and 45 Internet celebrities. I used an Instagram scraper (Tanupabrungsun, 2017a) and a Twitter scraper (Tanupabrungsun, 2017b) to collect their posts from 1/30/17 to 6/30/17, a 5-month period. I have also collected their daily follower counts since then using the Social Blade service<sup>7</sup>.

The final collection contains six-months of posts of each user along with their daily follower counts. For each post, the number of likes, comments (replies) and retweets were also recorded. In total, I have collected 132,823 posts from 78 celebrity accounts, consisting of 109,442 tweets and 23,381 Instagram posts. The table below presents the statistics of frequency per user, grouped by celebrity types and platforms.

---

<sup>2</sup> <https://www.forbes.com/sites/robertadams/2016/04/14/the-top-10-instagram-influencers/#144fd89a42ba>

<sup>3</sup> <http://www.elle.com/fashion/news/g25950/which-style-blogs-matter/>

<sup>4</sup> <http://www.marieclaire.com/fashion/a16668/fashion-bloggers/>

<sup>5</sup> <http://www.popcrunch.com/10-most-popular-non-celebrities-on-facebook/>

<sup>6</sup> <http://greatist.com/health/must-follow-health-and-fitness-twitter-accounts>

<sup>7</sup> <https://socialblade.com>

**Table 4.1** Statistics of frequency per user, grouped by celebrity types and platforms.

	Mainstream Celebrities				Internet Celebrities			
	Min.	Median	Mean	Max	Min.	Median	Mean	Max
<b>Twitter</b>	10.0	198.0	414.8	1926.0	20.0	191.0	559.3	3530.0
<b>Instagram</b>	13.0	218.0	383.4	2862.0	17.0	144.0	213.5	1421.0

### *B. Richness Annotations by AMT*

To collect the richness scores from the crowd, I implemented and distributed a web page for annotations on Amazon Mechanical Turk (AMT), the crowdsourcing service operated by Amazon. The implementation was customized specifically for the tasks, and consisted of two parts. The first part asked for background information about participants e.g., gender and age. The other part asked participants to rate eight unique posts (i.e., tweet or Instagram post) and two duplicate posts in order to measure *stability* using repeated measurements (Bland & Altman, 1986; Creswell, 1994). Specifically, the stability test is used to examine whether individuals varied their responses when the question was asked a second time over a short period of time. Responses from participants who give *unstable answers* would be dropped from an analysis. Below is an instruction for the tweet annotation task. Note that the instruction for Instagram annotation is almost identical except for the wording about Twitter.

*“Imagine the following tweet is from a celebrity you are following on Twitter e.g., pop stars, athletes or Internet celebrity. Think carefully about the way the tweet is constructed and answer if you agree with the following statements.”*

For each post, participants were asked to rate their agreement if the post matches the definitions of richness in identity, interaction and visibility separately i.e., either agree (rich) or disagree (lean). The three statement items were evaluated and pilot-tested to ensure they were understandable and measured what they intended to measure. This process was helpful to collect feedback on wording and clarity of the statements. The revised statements are presented below.

<i>Identity richness</i>	A post shows what the author is like; it shows his/her personality or character; it guides and controls the impression of readers on him/her; it gives the impression of his/her candid and uncensored looks.
<i>Interaction richness</i>	A post shows an attempt of the author to interact with followers, friends or peers; it reflects his/her attempt to maintain relationships with others; it creates a feeling that he/she is reachable.
<i>Visibility richness</i>	A post shows an attempt of the author to promote his/her presence beyond existing followers; it attempts to increase public exposure probably by bringing new audience to his/her account; it helps extend the reach of the post to a larger audience.

### *Datasets for Annotations*

To collect the labels, I drew samples of 1,000 Twitter and 1,000 Instagram posts from the larger collection of 132,823 presented earlier. The samples were stratified by users. Then, I created two AMT batches for Twitter and Instagram separately. Each batch consisted of 375 assignments, each of which consisted of ten posts (eight unique and two duplicate posts). To achieve high reliability, each post was annotated by three workers (Nowak & R ger, 2010). Both

AMT batches specified workers located in the US with an approval rate of greater than 95% to ensure high quality workers. I also instructed that they must be active Twitter or Instagram users in order to participate. Each approved assignment was rewarded with \$0.10. The following sections report the results of Twitter and Instagram labelling separately.

### *Twitter Annotations*

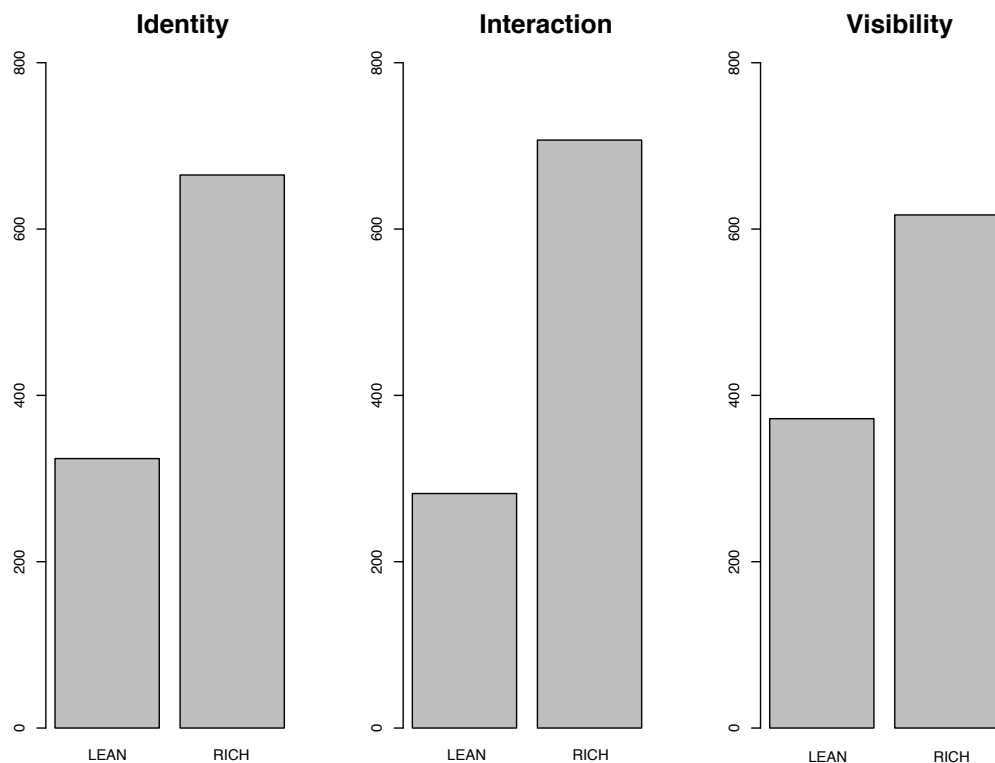
After 12 days, 375 unique workers completed the batch of Twitter annotations with an average of 5 minutes and 42 seconds per assignment. Amongst 375 users, only five users gave unstable responses (Bland & Altman, 1986; Creswell, 1994). In other words, they gave different answers to the same question. Thus, all of their annotations were removed from further analysis. The second batch was created to collect more responses from another five users, whose responses were stable.

The workers consist of 60.20% self-identified as female, 38.29% male and 1.5% did not wish to answer. A majority of the workers were between the age of 25-40 (53.65%), 26.19 % were younger than 25, 18.39% were 41-60 and 1.5% was older than 60. Most of the workers identified themselves as Caucasian (69.77%) followed by Asian (8.82%), African-American (8.31%), Hispanic (7.56%), Native American (1.26%) and Other (2.77%); the rest did not want to answer. For education, almost half of the workers had a college degree (49.87%, or a high school degree (29.22%), a graduate degree (19.14%) and the rest did not want to answer. When asked about their *Twitter-self*, 47.1% of workers identified themselves as a Lurker (rarely post, mostly read), 29.47% as a Retweeter/Liker (rarely post, mostly retweet/like others) and the rest



as a Poster (post frequently). The average number of accounts they were following on Instagram is 435.5 and the average followers they had was 435.

Before obtaining the final richness annotations, I calculated Krippendorff's alpha multi-coder agreement to compare the results from the three workers in each dimension (Krippendorff, 2012). The alpha coefficients of 0.71, 0.78 and 0.76 for identity, interaction, and visibility measures show that the annotations from different workers were reliable. Then, I used a *majority voting* technique to obtain the final annotations i.e., if the three workers annotated a post as rich, lean and rich, I labeled the tweet as rich. The distributions of final annotations are illustrated below. For all dimensions, the tweets tend to be annotated as rich more often i.e., the workers tend to see the tweets as rich in identity, interaction and visibility more often.



**Figure 4.1** Distributions of tweet annotations (Rich vs. Lean) by AMT workers, grouped by affordance richness dimensions. This shows that tweets tend to be labeled as *rich* more often.

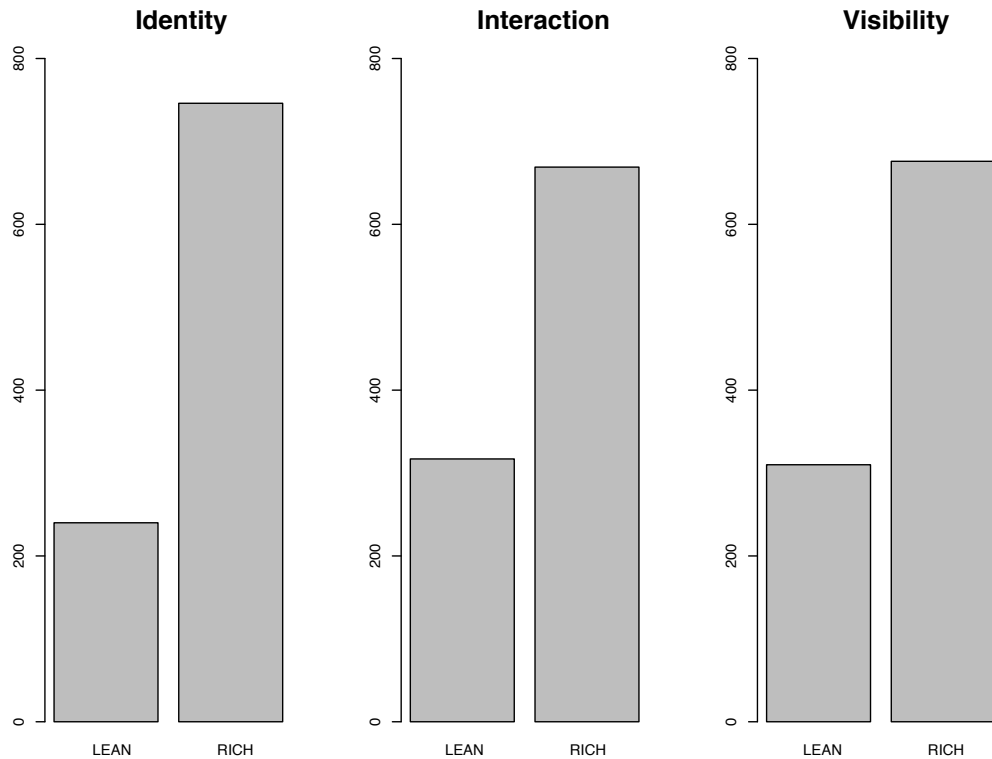
### *Instagram Annotations*

The Instagram batch was completed in 10 days by 375 unique workers with an average of 5 minutes and 56 seconds per assignment. Amongst 375 users, 11 users gave unstable responses and so their annotations were removed from further analysis. The second batch was created to collect more annotations from another 11 users, whose responses were all stable.

A majority of the workers identified themselves as female (73.84%), 25.64% as male and 0.51% did not wish to answer. 30.51% of them were younger than 25 years old, 55.89% were

between 25-40, 13.08% are 41-60 and 0.51% was older than 60. Most of the workers identified themselves as Caucasian (69.23%), African-American (11.54%), Asian (8.21%), Hispanic (7.69%), Other (2.05%) and 1% did not wish to identify. For education, a majority of the workers had a college degree (51.02%), a high school degree (32.82%) or a graduate degree (14.62%). When asked about their *Instagram-self*, 47.18% of workers identified themselves as a Liker/Commenter (i.e., rarely post, mostly like/comment others), 34.35% as a Poster (i.e., post frequently) and 18.46% as a Lurker (i.e., rarely post, mostly read). The average number of accounts they were following on Instagram is 283.48 and the average number of followers they had was 287.6.

As before, I calculated Krippendorff's alpha multi-coder agreement for each affordance dimension, comparing the results from the three workers (Krippendorff, 2012). The alpha coefficients of 0.74, 0.71 and 0.75 for identity, interaction, and visibility measure, show that the annotations from different workers were consistent. Then, I used a *majority voting* technique to obtain the final annotations e.g., if three workers annotated a post as rich, lean and rich, I labeled the post as rich. The distributions of final annotations are illustrated below. Similar to Twitter, the posts tend to be annotated as rich more often in all three dimensions of affordance richness.



**Figure 4.2** Distributions of Instagram post annotations by AMT workers, grouped by affordance richness dimensions. This shows that posts tend to be annotated as *rich* more often.

### *B. Modeling*

As shown above, the annotated datasets are unbalanced towards *rich* in all three dimensions of affordance richness. Imbalanced datasets can be problematic for standard classification algorithms as they tend to bias towards the majority classes and result in high misclassification rate for the minority classes (Estabrooks, Jo, & Japkowicz, 2004; Kotsiantis, Kanellopoulos, & Pintelas, 2006). As such, this study employs an ensemble learning technique rather than the standard learning algorithms to overcome the problem of imbalanced datasets.

## *Training Datasets*

The models were trained with the AMT annotated datasets. The targeting classes are annotations in three dimensions of affordance richness. The sizes of the training data are similar for both Twitter and Instagram. For each dataset, there are 1,200 instances, consisting of 1,000 annotations by AMT and 200 annotations by me. The additional 200 instances were added to the training datasets to improve the performance of the models. For both datasets, majority of the instances belong to the *rich* class.

**Table 4.2** presents the numbers of instances (%) labelled as *rich* in each dimension, for each platform.

	<b>Twitter</b>	<b>Instagram</b>
<b>Dimension</b>	<b>Annotated Data (%rich) n=1,200</b>	<b>Annotated Data (%rich) n=1,200</b>
Identity	61.25%	70.92%
Interaction	68.83%	53.00%
Visibility	60.17%	68.50%

## *Models Training*

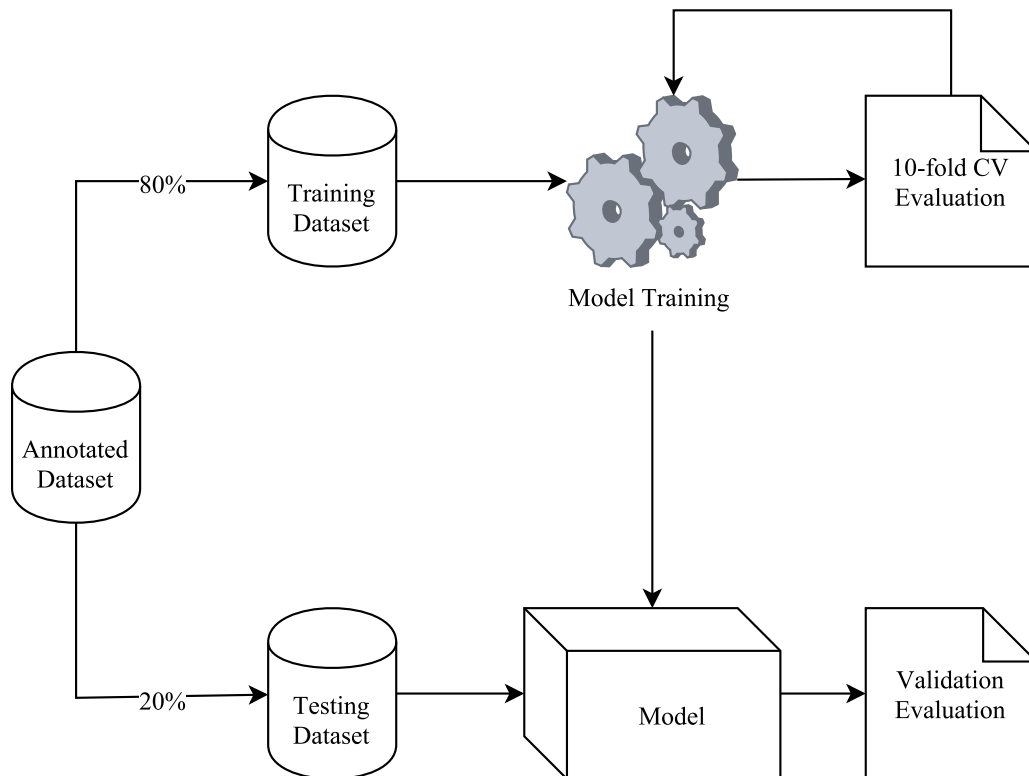
Given that the AMT annotations are dichotomous (rich or lean), I formulated the problems as a binary classification task. In an attempt to achieve high predictive power and overcome the problem of imbalanced datasets, I adopted an *ensemble learning* technique.

Ensemble learning is a technique for improving the predictive power of supervised algorithms by combining multiple *weak models* (i.e., low predictive power) to make a *stronger model* (i.e., high predictive power) (Opitz & Maclin, 1999; Rokach, 2010).

There are two common ensemble techniques: *bagging* and *boosting*. BAGGing or Bootstrap AGGregation builds an ensemble model by combining multiple models trained on different samples using bootstrap sampling. Specifically, each sample is drawn from the whole dataset with replacement. The technique is known for an ability to reduce variance i.e., making the model more generalizable to different datasets (Breiman, 1996). The Boosting technique constructs an ensemble model by incrementally training new models on the whole dataset, but instances might be weighed differently (Schapire, 1990). Specifically, a first model is trained with all training instances equally weighed. In the next iteration, a new model is trained with a focus on correcting the misclassified instances from the previous iteration. That is, the misclassified instances are given more weights to supervise the model to pay more attention in those instances. The training process is iterated until it satisfies the stopping criteria e.g., convergence of performance (no improvement of scores) and/or reaching the specified number of iterations.

One of the most popular Boosting algorithms is Adaptive Boosting with Decision Stump or *AdaBoost*. The algorithm was developed by Freund and Schapire (Freund & Schapire, 1995). It uses *decision trees* as weak learners and has been proven to overcome limitations of traditional algorithms by reducing both bias and variance in prediction (Ratkiewicz et al., 2011). Yet, some prior work has found that the algorithm could over-fit the data (X. Li, Wang, & Sung, 2005; Rätsch, Onoda, & Müller, 1998).

To prevent overfitting, I used an 80/20 hold-out test method to divide the annotated dataset into *training* and *testing* data i.e., 80% training and 20% testing. For AdaBoost, the only parameter that needed to be tuned was the number of trees, or the number of weak learners. I used a 10-fold cross validation to select the optimal parameter. Specifically, the training data was divided into 10 folds; a model was trained on nine folds and then tested on the other fold. The parameter and set of predictors that gave the highest average performance was then selected. The model training process is illustrated below.



**Figure 4.3** Model training process. The annotated dataset was split into training and testing datasets. The training dataset was used to identify an optimal parameter tuning using 10-fold cross validation. The testing dataset was used to evaluate the performance of the optimal model.

To evaluate the performance of the testing models, I report the *confusion matrix*, *precision*, *recall* and *F1-score* measures. The confusion matrix describes the performance of a supervised learning algorithm through a contingency table with two dimensions: actual and predictions. For a binary classification like my study, a confusion matrix consists of four entries. The first entry is *true negative* which includes *lean* messages correctly classified as *lean*. Second, *false negative* includes *rich* messages incorrectly classified as *lean*. Next, *false positive* includes *lean* messages incorrectly classified as *rich*. Lastly, *true positive* includes *rich* messages correctly classified as *rich*.

**Table 4.3** The entries of a confusion matrix for binary classifications.

		Actual	
		Lean	Rich
Prediction	Lean	True Negative: Lean as Lean	False Negative: Rich as Lean
	Rich	False Positive: Lean as Rich	True Positive: Rich as Rich

Precision measures the correctness of the classification (e.g., messages predicted as *rich* are indeed *rich* as per the annotations). A perfect precision score of one suggests that the classification judgements are credible or that the predictions are accurate. However, the precision measure does not show to what extent the models are able to detect relevant observations. This is where *recall* comes to play. Recall measures the ability of a model to include all relevant observations (e.g., all *rich* messages in the annotated data are classified as *rich*). A perfect recall of one suggests that a model has the ability to capture all relevant observations. The two measures are typically inversely related. For example, a model with high precision and low recall



suggests that its predictions are accurate (e.g., the messages predicted as *rich* are credible) but the model fails to include a lot of relevant observations (e.g., the model fails to predict a lot of *rich* messages as *rich*). The balance of these two measures depends on the context of a study. For example, a credit card fraud detection is typically in favor of high recall because the cost of missing a fraudulent transaction is higher than the cost of incorrectly identified a transaction as a fraud. For my study, both precision and recall are equally important and so I adopted a balanced measure, an F1-score, or the harmonic mean of precision and recall. An F1-score of one represents perfect precision and recall. Contra wise, a zero would indicate no correct classifications and no real observations were included.

To evaluate my models, I performed comparative analyses on a model's attribute values to characterize the misclassified data points in each entry of confusion matrices. Specifically, I looked at the central values of the model's attributes to examine if they exhibited similar or different patterns across the matrix e.g., whether or not the attribute values of the data points in *true positive* and *false positive* groups are different. For each entry of the confusion matrix, I reported *central values* of the attributes using the *mode* for the categorical attributes and the *median* for the numeric attributes. Note that some entries of my confusion matrices are small (i.e., less than 30) and potentially contain extreme values. Medians are generally more robust to skewed distributions and small datasets; thus, I chose to report the *median* rather than the *mean*. The following sections present discussions around the performance of each of the models separately.

As shown below, the misclassifications exhibit consistent patterns across the models. Specifically, the central values of the misclassified instances (i.e., median and mode) deviated from the correctly classified instances but they were closer to the other class. I note suggestions

on how to improve the models throughout the analyses and summarize at the end of this sections.

In total, I developed six classification models for each of the three affordance dimensions, for Twitter and Instagram. The targeting classes are annotations in each dimension of affordance richness. Table 4.4 summarizes the performances of all six classification models. Each model was compared against a baseline model using a majority classifier (i.e., simply predicts the majority class in the dataset). I reported the Kappa statistic, or a normalized accuracy score by the baseline. The Kappa coefficients range from -1 to 1 where the baseline accuracy is zero, negative scores mean a model performs worse than the baseline and scores above zero show an improvement over the baseline. My Kappa coefficients range from 0.290 to 0.645, suggesting the models perform better than the baseline.

**Table 4.4** Summary of models' accuracy compared to baseline models.

	Twitter			Instagram		
	Baseline Model	My Model	Kappa	Baseline Model	My Model	Kappa
<b>Identity</b>	0.613	0.725	0.290	0.709	0.796	0.298
<b>Interaction</b>	0.688	0.826	0.442	0.530	0.833	0.645
<b>Visibility</b>	0.602	0.804	0.508	0.685	0.763	0.246

For each model, the target class was a richness label i.e., *rich* or *lean*, and the predictors depended on the dimension of affordance richness being modeled and the platform. For example, the predictors of the Twitter interaction richness model are the number of @mentions, second

person pronouns and retweets. The table below summarizes the performances of all models. The F1-scores range from 73.83% to 83.28%, indicating that the models perform sufficiently well.

**Table 4.5** Performance of classification models, showing high predictive power with F1-scores all over 73%.

<b>Dimension</b>	<b>Twitter F1-scores</b>	<b>Instagram F1-scores</b>
Identity	77.86%	79.15%
Interaction	83.28%	80.19%
Visibility	76.62%	73.83%

### Twitter

This section presents a discussion around the development and performance of the Twitter models. For each model, I examined the *confusion matrix*, *precision*, *recall* and *F1-score* and delved deeper into each entry of the confusion matrix. For all models, the precision scores are higher than the recall scores. This means that we can be more confident in the predicted *rich* messages that they are actually *rich*, but slightly less confident with the predicted *lean* messages as some of them could be *rich*.

### *Identity Affordance*

The first dimension of affordance richness aims at tackling the needs of celebrities in identity construction. On social media, the identity work is a social act of positioning the self in relation to others by sharing information that reflects one's identity, or what impression they want others to have of them (Marwick, 2015b; Mavroudis & Milne, 2016; Page, 2012). Of course, the form of identity work varies by platforms and the artifacts they support. As a textual driven platform, constructing identity on Twitter is regarded as *written-into-being* (Marwick, 2015a). That is, social media users can express themselves through textual updates. As of the time of this study, Twitter, as a microblogging service, had a 140-character limit. Users could also provide more information by embedding a URL, which is regarded as an information resource (Bennett, Segerberg, & Walker, 2014), or including embedded content such as photos and videos. Looking at the content of tweets, Naaman et al. (2010) suggested 2 types of Twitter users: *Meformer* and *Informer*. Meformers are users the majority of whose tweets focus on the self, and are more personal in nature, while Informers refer to those the majority of whose tweets are more about the dissemination of informational content. As noted in the literature review section, none of the users has only one type of message (Naaman et al., 2010); therefore, none is an absolute Informer or Meformer. Thus, I only borrowed their message categories to examine tweet nature, rather than the nature of the users. Specifically, I suggest the Meformer messages would be more useful for helping users construct their identity. I operationalized the category of Meformer as the use of first person pronouns (i.e., I, me, my, mine, we, us, our, and ours).

After multiple rounds of modeling, the best performing model consists of three predictors: *text\_length*, *first\_person\_count* and *has\_url*. The first predictor measures the length of a tweet after removing special characters, @mentions, URL, hashtags and retweet artifacts

(i.e., RT @username:), *first\_person\_count* is the count of occurrences of the first-person pronouns and the other predictor, *has\_url*, is a Boolean indicating if a tweet contains URLs.

The identity model has an F1-score of 77.86% with precision and recall of 77.86%. That is, we can be 78% confident that the predicted *rich* messages are correct but the uses of the predicted *lean* messages must be cautious as some of them are, in fact, *rich*. Specifically, the model incorrectly predicted 22% of the *rich* messages as *lean* ( $100 - 77.86 = 22.14$ ).

**Table 4.6** Performance of the Twitter’s identity model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	86	31	77.86%	77.86%	77.86%
	Rich	31	109			

For this model, the attributes are *text\_length*, *first\_person\_count* and *has\_url*. For all but the *has\_url* attribute, the false negative group (i.e., the 22% of the *rich* messages that were predicted as *lean*) has lower median than the true positive group (*rich* messages predicted as *rich*). However, the values are closer to those of the true negative group (*lean* messages predicted as *lean*). The *has\_url* attribute, however, is similar to those in the true positive group. Looking closer at some tweets in the false negative group, I found that they seemed to make a commentary on an issue, share information or news. This suggests that the public could still perceive such tweets as *rich* in identity, or that the tweets could reflect one’s identity through the sharing of information without an explicit stancetaking. The model, however, could not detect this signal and so such *rich* messages were incorrectly classified as *lean*.

The messages in the false positive group (i.e., the 20% of the predicted *rich* that were incorrect) have similar *first\_person\_count* and *has\_url* values to those of the true positive group. Most of the tweets in the false positive group are relatively short and seem to be part of a bigger conversation. The public perceived such messages as *lean* in identity probably because they were not informative by themselves but could be more expressive within the context where they occurred. However, the model could not distinguish the content and so such *lean* messages were wrongly classified as *rich*.

In future work, we could improve the model by collecting more annotations for tweets that exhibit similar characteristics to the misclassified instances. We could also content analyze tweet texts to understand the underlying nature e.g., whether they are conversational tweets or complete by themselves. I note this is beyond the scope of my study but opens a direction for future work.

**Table 4.7** Attribute values of the Twitter's identity model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>text_length</i>	60.00	44.00
	<i>first_person_count</i>	0.00	0.00
	<i>has_url</i>	False	True
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>text_length</i>	55.00	88.00
	<i>first_person_count</i>	1.00	1.00
	<i>has_url</i>	True	True

### *Interaction Affordance*

This dimension of the framework addresses the needs to interact with fans. On social media, celebrities treat their followers as an aggregated fan base. Interaction develops and maintains audience through responding, or reaching out to fans, or creating an illusion of such activities. On Twitter, users can interact with others through addressivity mechanisms i.e., @mention and @reply (Honeycutt & Herring, 2009), as well as a public recognition feature like retweets (boyd et al., 2010; Marwick & boyd, 2011; Metaxas et al., 2015; Pennington et al., 2016) to show that they are listening or acknowledge the original author. Users can address the audience by using second person pronouns such as *you* and *guys* (Quercia et al., 2011; Raun,

2018). The interaction richness of a tweet is the uses of @mentions, @replies, retweets and second-person pronouns.

The best performing model consists of three predictors: *mentions\_count* (@replies are included), *second\_person\_count* and *is\_retweet*. The first predictor is the number of @username contained in a tweet, *second\_person\_count* is the count of occurrences of the second-person pronouns (i.e., you, your, y'all, guys and folks) and the other predictor, *is\_retweet*, is a Boolean indicating if a tweet is a retweet.

The interaction model has an F1-score of 83.28% with precision of 81.88% and recall of 84.72%. That is, the predicted *rich* messages are correct 82% of the time but 15% of the *rich* messages could not be identified correctly. Specifically, the model incorrectly predicted 15% of the *rich* messages as *lean* ( $100 - 84.72 = 15.28$ ).

**Table 4.8** Performance of the Twitter's interaction model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	69	22	81.88%	84.72%	83.28%
	Rich	27	122			

For this model, the attributes are *mentions\_count*, *is\_retweet* and *second\_person\_count*. For *mentions\_count*, the false negative group (the 15% of the *rich* messages that were misclassified as *lean*) has a lower median than the true positive group (*rich* messages predicted as *rich*) but identical with those of the true negative group (*lean* messages predicted as *lean*).



However, *is\_rt* exhibits a different pattern. Specifically, it is *True* (i.e., a tweet is a retweet) for the true positive and false negative groups. That is, the public tended to perceive a retweet as *rich* in interaction but the model failed to identify some of these retweets as *rich*. This suggests that *mentions\_count* is more discriminative than *is\_rt* for the classification of this dimension.

One way to improve the model is to collect more annotations particularly for *retweets*.

Specifically, we could construct a sample of retweets and ask AMT workers to annotate the interactional richness label. I expect the additional annotated retweets to provide more insights into this observation and helpful for the model to correctly classify retweets.

The other misclassified instances are those in the false positive group (*lean* messages predicted as *rich*), their *mention\_counts* values are closer to those of the true positive group. This suggests that the public perceived some tweets as *lean* in interaction even though they had @mentions embedded. A closer look at the tweets in this misclassified group suggested that some of the tweets were product endorsement and the @mentioned accounts were not a person, but they were brands or organizations. However, the model could not distinguish the @mention accounts and classified such *lean* messages as *rich*. We could potentially improve the model if we have an exhaustive list of brands or content analyze the intention of the poster (e.g., whether they are product endorsement).

**Table 4.9** Attribute values of the Twitter's interaction model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>mentions_count</i>	0.00	0.00
	<i>is_rt</i>	False	True
	<i>second_person_count</i>	0.00	0.00
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>mentions_count</i>	1.00	1.00
	<i>is_rt</i>	False	True
	<i>second_person_count</i>	0.00	0.00

### *Visibility Affordance*

The last dimension of richness tackles the need to promote visibility. As noted earlier, attention is a scarce resource in social media. A piece of information needs to compete against others to become visible (Brighenti, 2010). Visibility work promotes content and competes for public attention. Twitter supports a mechanism to increase visibility beyond the existing follower network. The use of hashtags is widely recognized as increasing public exposure by making the posts appear in the search feature. When hashtags are used by a large number of users, they become a trending topic on Twitter. The literature also suggested that public attention is time dependent. For example, an interview with Internet celebrities in Singapore indicated that the best times to get likes were from 8-10am and 7-9pm weekdays (Abidin, 2014).

The visibility model consists of three predictors: *hashtags\_count*, *week\_day* and *time*. The first predictor is the count of hashtag occurrences, *week\_day* is a Boolean indicating if a post was created during the week (i.e., Monday to Friday) or weekend (*True* if week days otherwise *False*), and *time* is discretized into a categorical variable: Morning (12am-9am), Afternoon (9am-6pm) and Night (6pm-12am). Note that this discretization was based on the standard working time; thus, Morning refers to the before work hours and Night refers to the after work hours.

The visibility model has an F1-score of 76.62% with precision (82.80%) slightly higher than recall (71.30%). That is, the predicted *rich* messages are correct 83% of the time and there are 29% of the *rich* messages that cannot be correctly identified ( $100 - 71.30 = 28.70$ ).

**Table 4.10** Performance of the Twitter's visibility model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	116	31	82.80%	71.30%	76.62%
	Rich	16	77			

For this model, the attributes are *hashtags\_count*, *week\_day* and *time*. For the *hashtags\_count* attribute, the median of the false negative group (the 29% of the *rich* messages that were misclassified as *lean*) is close to those of the true negative group (*lean* messages predicted as *lean*). However, *week\_day* values are similar to those of the true positive group as *False*. In other words, they were tweeted during the weekend. Looking at the false positive group (*lean* messages predicted as *rich*), their *hashtags\_count* values are similar to those of the true

positive group but their *week\_day* values are similar to those of the true negative group.

Together, this suggests that *hashtags\_count* is the most discriminative attribute for the classification of this dimension.

In future work, we could potentially improve the model by collecting more annotations for tweets which exhibit the characteristics of the misclassified instances. I expect the additional training data to be helpful for the model to learn the patterns and correctly classify *rich* tweets with low *hashtags\_count* or *lean* tweet with high *hashtags\_count*.

**Table 4.11** Attribute values of the Twitter’s visibility model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>hashtags_count</i>	0.00	0.00
	<i>week_day</i>	True	False
	<i>time</i>	Morning	Afternoon
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>hashtags_count</i>	1.00	1.00
	<i>week_day</i>	True	False
	<i>time</i>	Afternoon	Afternoon

## *Instagram*

As before, I report the *confusion matrix*, *precision*, *recall* and *F1-score* for each of the Instagram models, followed by a discussion on each entry of the confusion matrix. Similar to Twitter's models, all precision scores are higher than the recall scores, suggesting that we can be more confident in using the predicted *rich* messages but should be cautious that some of the actual *rich* messages are misclassified as *lean*.

## *Identity Affordance*

Instagram is a visually driven platform. It provides opportunities to construct identity through a visual medium along with caption text. Previous work has found different ways people post images on Instagram, such as captioned images (i.e., an image of text or a quote), selfies and non-human images (Hu et al., 2014). Bakshi et al. (2014) found that images with the presence of faces engaged audience better. To conform with the definition of microcelebrity as turning oneself into media content, I measured identity richness as the presence of faces by using the face detection module provided by Google Cloud Vision API<sup>8</sup>. This service is very robust with an accuracy of 99.63% (Schroff, Kalenichenko, & Philbin, 2015). I also used caption text length and the presence of a URL.

After multiple rounds of modeling, I select a model of three predictors: *faces\_count*, *first\_person\_count* and *text\_length*. The first variable is a numerical value for the number of human faces presented in a photo as detected by Google API, *first\_person\_count* is the count of

---

<sup>8</sup> <https://cloud.google.com/vision/docs/>

occurrences of the first-person pronouns (i.e., I, me, my, mine, we, us, our, and ours) and the other variable, *text\_length*, is the length of caption text after removing special characters, @mention, URL and hashtags.

The identity model has an F1-score of 79.58% with precision (86.11%) higher than recall (73.22%). That is, we can be 86% confident that the predicted *rich* messages are correct but be cautious that 27% of the *rich* messages could not be correctly identified ( $100 - 73.22 = 26.78$ ).

**Table 4.12** Performance of the Instagram’s identity model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	98	34	86.11%	73.22%	79.58%
	Rich	15	93			

For this model, the attributes are *faces\_count*, *first\_person\_count* and *text\_length*. For all attributes, the false negative group (the 27% of the *rich* messages that were misclassified as *lean*) has closer median values to the true negative group (*lean* messages predicted as *lean*) than true positive group (*rich* messages predicted as *rich*). Many of the posts in this misclassified group were descriptive e.g., about holidays or meal preparations. While the public perceived such posts as *rich*, the model could not identify this characteristic of the posts and incorrectly classified them as *lean*.

The messages in the false positive group (*lean* messages predicted as *rich*) also have similar *faces\_count* and *text\_length* values to those of the false negative and true negative groups. However, their *first\_person\_count* values are closer to the true positive group. Some of

the posts in this misclassified group seemed to be promotional posts e.g., product endorsement. Obviously, the public perceived such posts as *lean* in expressing one's identity but the model could not distinguish the content and so such *lean* messages were wrongly classified as *rich*. This also suggests that the *first\_person\_count* is more relevant for the classifications of the *rich* class.

In future work, we could conduct a content analysis on caption texts to identify the topics of the post and add them to the list of attributes to help the model better understand the characteristics of the posts.

**Table 4.13** Attribute values of the Instagram's identity model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>faces_count</i>	0.00	0.00
	<i>text_length</i>	24.00	17.00
	<i>first_person_count</i>	0.00	0.00
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>faces_count</i>	0.00	1.00
	<i>text_length</i>	17.00	44.00
	<i>first_person_count</i>	1.00	1.00

### *Interaction Affordance*

Instagram is quite short on interactivity. The platform does not have a conversational mechanism. Although users can interact by exchanging comments and @mention others, the comment section shows only the latest few messages. Some highly engaged practitioners frequently respond to fans' comments (Marwick, 2015a), or by tagging other users in their posts (Ward, 2016). The interaction richness of an Instagram post is operationalized as the uses of @mentions and second person pronouns.

The best performing model uses two predictors: *mentions\_count* and *second\_person\_count*. The first predictor is a numerical value of the number of @usernames contained in a post. The other variable is also a numerical value of the occurrences of second-person pronouns (i.e., you, your, y'all, guys and folks).

The interaction model has an F1-score of 80.19% with precision (94.44%) higher than recall (69.67%). That is, we can be 94% confident that the predicted *rich* messages are correct but some of the *rich* messages were misclassified as *lean*. Specifically, the model incorrectly predicted 30% of the *rich* messages as *lean* ( $100 - 69.67 = 30.33$ ).

**Table 4.14** Performance of the Instagram's interaction model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	113	37	94.44%	69.67%	80.19%
	Rich	5	85			



For this model, the attributes are *mentions\_count* and *second\_person\_count*. For the *mentions\_count* attribute, the false negative group (the 30% of the *rich* messages that were predicted as *lean*) has lower median than the true positive group (*rich* messages predicted as *rich*) but identical with those of the true negative group (*lean* messages predicted as *lean*). Their *second\_person\_count* attributes, however, are closer to the true positive group. Therefore, this suggests that *mentions\_count* is more relevant for the classifications of the *rich* class. We could potentially improve the model by collecting more annotations particularly for the posts that exhibit this characteristic (i.e., low *mentions\_count* but high *second\_person\_count*).

The messages in the false positive group (*lean* messages predicted as *rich*) also have similar *second\_person\_count* values to those of the true negative. However, their *mentions\_count* values are closer to the true positive group. Some of the misclassified posts in this group are product endorsement and the @mention accounts are mostly brands, magazines and organizations. While the model could not distinguish the accounts, the public could and so perceived such messages as *lean* in interactions. We could potentially improve the model if we have an exhaustive list of brands, and by content analyzing the texts to identify if they are promotional posts or endorsing a product.

**Table 4.15** Attribute values of the Instagram’s interaction model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>mentions_count</i>	0.00	0.00
	<i>second_person_count</i>	0.00	1.00
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>mentions_count</i>	2.00	2.00
	<i>second_person_count</i>	0.00	1.00

### *Visibility Affordance*

Similar to Twitter, Instagram affords a mechanism to increase visibility beyond the existing follower network. The use of hashtags is widely recognized as increasing public exposure by making the posts appear in the search feature. When hashtags are used by a large number of users, the posts could get featured on the Explore tab. Another factor is date and time. The literature suggests public attention is time dependent and there exist some particular times of the day when the traffic might be heavier, and so celebrities tend to get more responses from their fans (Abidin, 2014).

The visibility model consists of three predictors: *hashtags\_count*, *week\_day* and *time*. The first predictor is the count of hashtag occurrences, *week\_day* is a Boolean variable (*True* for weekdays otherwise *False*) and *time* is discretized into a categorical variable: Morning (12am-9am), Afternoon (9am-6pm) and Night (6pm-12am).

The visibility model has an F1-score of 73.83% with precision (83.16%) higher than recall (66.39%). That is, we can be 83% confident that the predicted *rich* messages are correct but be cautious that 34% of the *rich* messages could not be correctly identified ( $100 - 66.39 = 33.61$ ).

**Table 4.16** Performance of the Instagram’s visibility model.

		Actual		Precision	Recall	F1-score
		Lean	Rich			
Prediction	Lean	105	40	83.16%	66.39%	73.83%
	Rich	16	79			

For this model, the attributes are *hashtags\_count*, *week\_day* and *time*. For the *hashtags\_count* attribute, the median of the false negative group (the 34% of *rich* messages that were predicted as *lean*) are closer to those of the true negative group (*lean* messages predicted as *lean*). The mode of *week\_day* attribute, however, is similar to the true positive group. The other misclassified group is false positive (*lean* messages predicted as *rich*), their *hashtags\_count* values are closer to those of the true positive group although their *week\_day* value is more similar to the true negative group. Together, this suggests that *hashtags\_count* is the most discriminative attribute for the classification of this dimension.

Similar to the Twitter’s visibility model, we could potentially improve the model by collecting more annotations for tweets which exhibit the characteristics of the misclassified instances. I expect the additional training data to be helpful for the model to learn the patterns

and correctly classify *rich* posts with low *hashtags\_count* or *lean* tweet with high *hashtags\_count*.

**Table 4.17** Attribute values of the Instagram’s visibility model.

		<b>Actual=Lean (Median/Mode)</b>	<b>Actual=Rich (Median/Mode)</b>
<b>Prediction=Lean</b>		<i>True Negative</i>	<i>False Negative</i>
	<i>hashtags_count</i>	0.00	0.00
	<i>week_day</i>	False	True
	<i>time</i>	Afternoon	Afternoon
<b>Prediction=Rich</b>		<i>False Positive</i>	<i>True Positive</i>
	<i>hashtags_count</i>	1.00	2.00
	<i>week_day</i>	False	True
	<i>time</i>	Afternoon	Afternoon

In summary, the analyses of the attribute values show consistent patterns. The central values of the misclassified instances (e.g., median and mode) deviated from the instances in the correctly classified instances but they were closer to the other class. In future work, we could improve the models by collecting more annotations particular for the posts that exhibit the characteristics identified earlier i.e., the posts that have attribute values similar to the misclassified instances. The additional training data are expected to help the models learn the misclassified patterns and improve their performance. Alternatively, we could employ an active learning approach, a semi-supervised learning to iteratively add more training data based on

some criteria e.g., *uncertainty* or choosing the instances which a model has the least confidence in prediction. We could also develop *content features* by conducting content analysis on the text data to identify the intention of the poster or underlying nature of the posts (e.g., the topical focus of the text), and add the content features to the attributes of the classification models. Although adding more attributes to the models would be useful for handling the misclassifications, they could lead to *overspecified models*. Such models are likely to over-fit the training data and do not generalize well to the unseen data.

### *C. Predictions*

With the identified optimal parameter setting, I re-trained the models on the full training data and used them to predict the richness scores of the larger unlabeled datasets. The table below compares the proportions of instances annotated as rich, in the training data vs. the predictions. As shown here, the distributions are very similar thus suggesting the predictions are reliable.

**Table 4.18** Distributions of training data and predictions, showing that they are very close with the differences being less than 5.12%.

	<b>Dataset</b>	<b>Training Data (% rich) n=1,200</b>	<b>Prediction (% rich)</b>	<b>% Difference (training - prediction)</b>
<b>Twitter</b> <b>n=109,442</b>	Identity	61.25%	56.13%	5.12%
	Interaction	68.83%	63.73%	5.10%
	Visibility	60.17%	55.93%	4.24%
<b>Instagram</b> <b>n=23,381</b>	Identity	70.92%	72.56%	-1.64%
	Interaction	53.00%	50.06%	2.94%
	Visibility	68.50%	66.77%	1.73%

For the predictions, the magnitudes of correlation coefficients between each pair of richness dimensions are between 0.11 to 0.18 for Instagram and 0.001 and 0.007 for Twitter data. This suggests that the richness measures are not correlated and so capture different signals of affordance richness.

**Table 4.19** Correlation matrices of richness scores. The coefficient magnitudes are all lower than 0.2, suggesting that the richness measures are not correlated.

	<b>Twitter</b>				<b>Instagram</b>		
	Identity	Interaction	Visibility		Identity	Interaction	Visibility
Identity	1.000	-0.114	-0.119	Identity	1.000	-0.004	0.001
Interaction	-0.114	1.000	0.179	Interaction	-0.004	1.000	-0.007
Visibility	-0.119	0.179	1.000	Visibility	0.001	-0.007	1.000

## 4.2. Score Analysis

To examine the richness scores, I performed statistical tests such as the *t*-test for means equality, and constructed linear regression models using richness predictions as the independent variable. The dependent variables are audience responses using two proxies: followership and audience engagement. The datasets and analyses are explained below.

### 4.2.1 Dataset

For the following analyses, the dataset is the collection of tweets and Instagram posts from both mainstream and Internet celebrities presented earlier. For each post, the richness scores of the three affordance dimensions were predicted by the classification models. Finding the scores for each post allowed me to aggregate scores in different ways. As shown below, some analyses were at the post level, and some were at the user level. The table below presents the

proportion of posts categorized as rich in each affordance dimension, grouped by celebrity types for each platform.

**Table 4.20** Proportions of posts categorized as rich in each affordance dimension, grouped by celebrity types, for each platform.

	Twitter (n=109,442)				Instagram (n=23,381)		
	Identity	Interaction	Visibility		Identity	Interaction	Visibility
<b>Internet</b> <b>(n=25,170)</b>	13,757 (54.66%)	16,780 (66.67%)	16,665 (66.21%)	<b>Internet</b> <b>(n=10,581)</b>	7,127 (67.35%)	5,436 (51.38%)	7,277 (68.77%)
<b>Mainstream</b> <b>(n=84,272)</b>	51,942 (61.64%)	51,884 (61.57%)	55,554 (65.92%)	<b>Mainstream</b> <b>(n=12,800)</b>	8,983 (70.18%)	7,464 (57.53%)	7,152 (55.88%)

#### 4.2.2 Comparisons by platforms

The first question asks if the celebrity practices differ on different platforms: ***RQ1**: Along the core practices, how do celebrities engage in different activities on different social media platforms?* The analysis was thus performed at the user level. That is, the richness predictions were aggregated by users where each user was represented by the ratio of rich labels over all labels.

To answer the question, I conducted a series of paired *t*-tests to examine the similarity or difference of richness scores in each richness dimension, where each pair represented a user on



Twitter and Instagram. The tests are generally used to determine whether the mean difference between the paired samples is zero with the null hypothesis being that the true mean difference is zero. Before conducting the tests, I performed and reported a normality test to ensure the assumptions are satisfied i.e., the differences between pairs are independently and normally distributed without outliers, although the tests are generally robust (Montgomery, 2017).

#### *4.2.3 Comparison by celebrity types*

The second research question: **RQ2:** *How are the practices similar and/or different amongst mainstream and Internet celebrities engaging in microcelebrity?* To examine the relationship between richness scores and celebrity types, i.e., mainstream and Internet famous, I conducted another series of *t*-tests on each of the richness dimensions at the group level. Note that the first analysis was at the user level and so the scores are aggregated by users. This analysis, however, was at the group level and so the scores were aggregated by celebrity types. The assumptions of regular *t*-tests are that the samples are independently and normally distributed and that samples are large enough (i.e., 30 samples are generally acceptable) (Montgomery, 2017). As such, I performed and reported another normality test to ensure the assumptions were satisfied.

#### *4.2.4 Audience response*

The next two research questions inquire into how the audience responds to microcelebrity practices: **RQ3:** *How do the audiences respond to different types of strategies*

when controlled for celebrity types? and **RQ4**: How do the audiences respond to the changes in microcelebrity strategies? As noted earlier in the literature review, I used two proxies to measure audience's response: engagement and growth. The engagement scores are calculated using the equations derived from the Facebook's (2017) official engagement formulas.

$$\text{Equation 4.2} \quad \text{Facebook Engagement} = \frac{\# \text{ Likes} + \# \text{ Comments} + \# \text{ Share}}{\text{Post Reach}}$$

For Instagram, the numerator is the summation of numbers of *likes* and *comments*, divided by *number of followers*.

$$\text{Equation 4.3} \quad \text{Instagram Engagement} = \frac{\# \text{ Likes} + \# \text{ Comments}}{\# \text{ Followers}}$$

For Twitter, the numerator is the summation of numbers of *favorites*, *replies* and *retweets*, and divided by *number of followers*.

$$\text{Equation 4.4} \quad \text{Twitter Engagement} = \frac{\# \text{ Favorites} + \# \text{ Replies} + \# \text{ Retweets}}{\# \text{ Followers}}$$

It is important to note that using single measurement units like Equations 4.2 - 4.4 allow for cross-platform analyses. In this way, I can examine the relationships between microcelebrity practices and an audience's engagement on Twitter and Instagram altogether. Otherwise, I would need to create a mapping of the forms on engagement on Twitter and Instagram (e.g., Twitter's *favorites* and Instagram's *likes*). This is essential because the mapping can be challenging as users on different platforms tend to behave differently – for example, while *like* is the primary form of engagement on Instagram, *favorite* is scarcely used on Twitter.

To examine the relationships between engagement scores and richness scores, I developed linear regression models using engagement scores as the dependent variables. The

independent variables were richness scores in three dimensions. I also controlled for the effects of the platforms and celebrity types. Note that the numbers of followers did not need to be controlled as they were already accounted for when calculating the engagement scores, more discussion on this is presented in the result section (Chapter 5).

Another set of models looked at the relationships between richness scores and the other measure of audience's response, changes in followers. As before, the independent variables were richness scores in three dimensions, the platforms and celebrity types. I also controlled for the numbers of followers on the first day of data collection.

To ensure the assumptions of linear regression were satisfied, I performed a residual analysis to examine if the residuals were normally distributed with a constant variance (Faraway, 2004). The normality assumption was validated with the Shapiro-Wilk (SW) test with the null hypotheses of normal distribution. The constant variance assumption was validated with the Non-Constant Variance (NCV) test with the null hypothesis of constant variance. I also performed a multicollinearity test and reported the Variance Inflation Factor (VIF) coefficients to ensure the models did not suffer from multicollinearity (Faraway, 2004; James, Witten, Hastie, & Tibshirani, 2013).

### **4.3 Summary**

In this chapter, I presented the design and development of the richness framework. With the established framework, the affordance richness scores of the larger dataset were predicted and could be adopted to study microcelebrity practices in numerous of ways. To answer my

research questions, I designed and conducted the analyses using statistical approaches such as *t*-tests and regression modeling.

In the next chapter I report the results of the quantitative analyses and introduce how the findings were used to inform the design of the follow-up qualitative study.

## CHAPTER 5

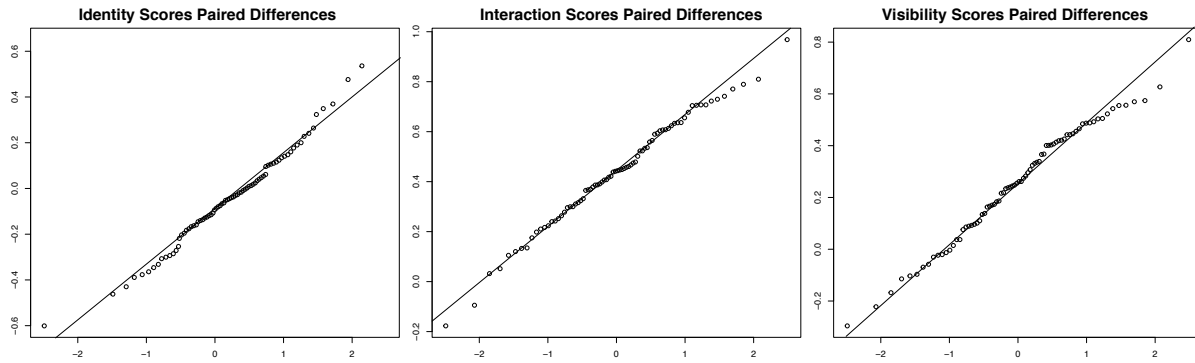
### QUANTITATIVE RESULTS AND FINDINGS

Following the analytical plans outlined earlier, I performed series of  $t$ -tests, and constructed linear regression models to explore the relationships between audiences' responses (i.e., engagement scores and changes in followers) and the richness measures. The regression models also included control variables as appropriate. The results are presented below.

#### 5.1 Comparisons by platforms

The first question asks if the celebrity practices differ on different platforms: *RQ1: Along the core practices, how do celebrities engage in different activities on different social media platforms?* The analysis was thus performed at the user level. That is, the richness predictions were aggregated by users where each user was represented by the ratio of rich labels over all labels.

I conducted a series of paired  $t$ -tests to examine the similarity of richness scores on different platforms. Each pair represented the richness scores of a user on Twitter and Instagram. As noted earlier, the assumptions of paired  $t$ -tests are that the differences between pairs are normally and independently distributed (Montgomery, 2017). The SW normality tests and Figure 5.1 show that the pair differences are normally distributed.



**Figure 5.1** Normal plots of the differences between pairs in each richness dimension. These plots show that the data are normally distributed and so, the assumption of paired  $t$ -tests is satisfied.

Table 5.1 presents the  $t$ -test results and Cohen's  $d$ , a measure of effect sizes. An effect is *negligible* when the magnitude is smaller than 0.2 (Cohen, 1992). The  $p$ -values are all lower than 0.01 and all magnitudes of Cohen's  $d$  are over 0.2, indicating strong evidence of differences in scores on Twitter and Instagram. For all but the identity dimension, the richness scores on Twitter are higher than those on Instagram. Therefore, I answer the first research question that the practices are different by platforms. Specifically, celebrities in my dataset use Instagram for constructing and/or expressing their identity more often than on Twitter. On the other hand, they use Twitter for interacting with others and promoting their visibility more often.

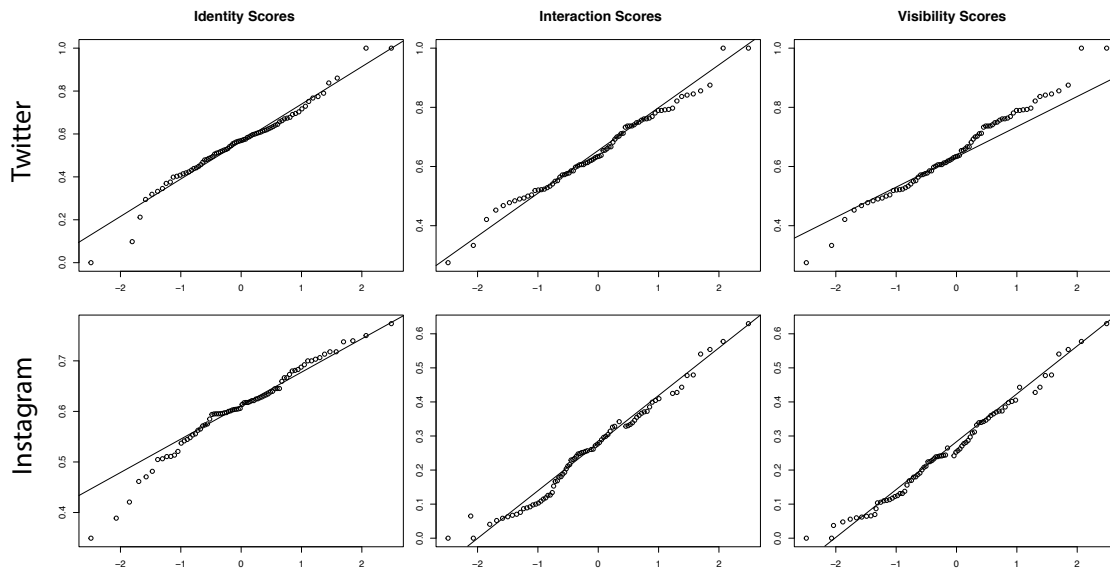
**Table 5.1** Statistical tests on the differences of scores by platforms. Note that the mean differences are the mean scores of Twitter subtracted by the mean scores of Instagram.

<b>Dimension</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>	<b>Cohen's <i>d</i> Effect Size</b>	<b>Mean Diff.</b>
Identity	-2.77	<0.01*	-0.314*	-0.05
Interaction	17.72	<0.01*	2.00*	0.44
Visibility	9.99	<0.01*	1.132*	0.25

## 5.2 Comparison by celebrity types

This analysis aims at answering the second research question: *RQ2: How are the practices similar and/or different amongst mainstream and Internet celebrities engaging in microcelebrity?* To examine the relationship between richness scores and celebrity types i.e., mainstream and Internet famous, I conducted another series of *t*-tests on each of the richness dimensions. As shown earlier, the practices differ by platforms. The tests were conducted on each platform separately to prevent the confounding effects of platform differences.

Before conducting the test, I ensured the assumptions of the tests were satisfied. With 33 samples in a group of mainstream celebrities and 45 samples in the other group of Internet celebrities, the assumption of large sample size is satisfied i.e., 30 samples are generally recognized as large enough (Montgomery, 2017). The SW normality tests and Figure 5.2 show that the samples are normally distributed.



**Figure 5.2** Normal plots of the samples in each richness dimension for each platform. These plots show that the data are normally distributed and so, an assumption of  $t$ -tests is satisfied.

The results in Table 5.2 show that only interaction richness on Instagram differs by celebrity types with the  $p$ -value less than 0.01 and Cohen's  $d$  of -0.508. That is, mainstream and Internet celebrities use the platforms similarly in all dimensions except for the interaction measure on Instagram where the Internet celebrities have significantly higher ratio of interaction-rich posts than the mainstream celebrities do. Therefore, I answer the second research question that only the interaction practices on Instagram are different by celebrity types. Specifically, the Internet celebrities in my dataset use Instagram for interacting with others more often than the mainstream celebrities do.



**Table 5.2** Statistical tests on differences of scores by the celebrity types for each platform. Note that the mean differences are the mean scores of the mainstream celebrities subtracted by the mean scores of the Internet celebrities.

	Twitter				Instagram			
Dimension	<i>t-val</i>	<i>p-val</i>	Cohen's <i>d</i>	Mean Diff.	<i>t-val</i>	<i>p-val</i>	Cohen's <i>d</i>	Mean Diff.
Identity	-0.229	0.819	0.048	0.007	-1.631	0.109	-0.178	-0.031
Interaction	-0.029	0.977	0.006	0.001	-2.732	<0.01*	-0.508*	-0.086
Visibility	0.468	0.642	-0.104	-0.014	0.303	0.763	0.065	0.011

### 5.3 Audience response

The next two research questions inquire about the responses from the public: **RQ3:** *How do the audiences respond to different types of strategies when controlled for celebrity types?* and **RQ4:** *How do the audiences respond to the changes in microcelebrity strategies?* As noted earlier, I used two proxies to measure an audience's response: *engagement* and *growth*. The engagement scores are calculated using the equations derived from the Facebook's (2017) official engagement formulas (Equation 4.2). For Instagram, the numerator is the summation of numbers of *likes* and *comments*, divided by number of *followers* (Equation 4.3). For Twitter, the numerator is the summation of numbers of *favorites*, *replies* and *retweets*, and divided by number of *followers* (Equation 4.4). The other measure of audiences' response, *audience growth* is operationalized as changes in numbers of followers which is calculated as the difference of numbers of followers recorded on the first and last day of data collection.

To examine the relationships between audiences' responses (i.e., engagement and growth) and my richness measures, I developed four sets of linear regression models. Through successive modeling process, I arrived at four high explanatory power models which satisfied the assumptions of linear regression (e.g., residuals are normally distributed with constant variance). In the sections below I discuss the iterative development of each regression model.

The first model explains the relationships between the engagement scores and the richness measures. The second model explains the relationships between the mean engagement scores and the *changes* of richness scores over time (i.e., within-user variance). Both of these models controlled for the effects of platforms and celebrity types. Note that the numbers of followers did not need to be controlled for as they were already accounted for when calculating the engagement scores. One particular avenue of model exploration I want to highlight is that I also constructed regression models that used the number of followers as a control variable instead of as part of the denominator of the engagement equations. While the significance and directions of the richness variables were similar, the models performed better, as per the *R*-Squared measures and residual analysis, when the number of followers was part of the denominator.

The other two models looked at the other measure of audiences' response: *audience growth*, operationalized as changes in numbers of followers. The first model explains the relationships between the changes in numbers of followers and the mean richness scores. The last model explains the relationships between the changes in numbers of followers and the *changes* of richness scores over time (i.e., within-user variance). Both of these models controlled for the effects of number of followers, platforms and celebrity types.

### ***Model 1: Engagement scores and richness measures***

To examine the relationships between engagement scores and richness scores, the first linear regression model was constructed at the post level. Each observation represents a post and consists of the engagement score, three richness scores, number of followers when the post was created, the platform where the post occurs and celebrity type of the poster. The header for the dataset of 132,823 records is:

Engagement score	Identity score	Interaction score	Visibility score	Number of followers	Platform	Celebrity type
------------------	----------------	-------------------	------------------	---------------------	----------	----------------

Before obtaining the final model, I experimented with different versions of modeling. The first version used the number of followers as a control variable, and calculated the engagement scores as the summation of the number of *likes* and *comments* for Instagram posts; and the summation of the number of *likes*, *replies* and *retweets* for Twitter data. I constructed the model using the raw engagement scores (i.e., no transformation) as the dependent variable. The independent variables were the richness scores in three dimensions. The model controlled for the effects of number of followers, platforms and celebrity types. Note that all continuous variables were normalized to make the interpretations easier.

*EngagementScore:* The continuous dependent variable. This variable was transformed using a logarithm function.

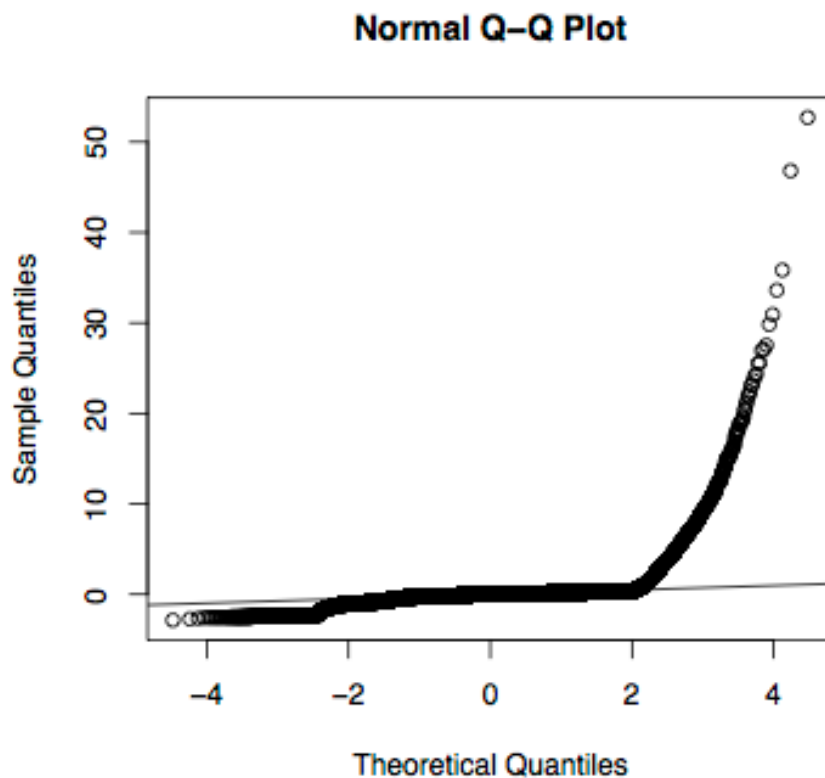
*Identity:* A Boolean variable, 1 if a post was categorized as rich in identity measure otherwise 0.

<i>Interaction:</i>	A Boolean variable, 1 if a post was categorized as rich in interaction measure otherwise 0.
<i>Visibility:</i>	A Boolean variable, 1 if a post was categorized as rich in visibility measure otherwise 0.
<i>Followers</i>	A normalized continuous variable represents the number of followers when a post was created.
<i>CelebrityType:</i>	A Boolean variable with mainstream as a base state i.e., 1 if an account was a mainstream celebrity otherwise 0.
<i>Platform:</i>	A Boolean variable with Twitter as a base state i.e., 1 if a post was a tweet otherwise 0.

To ensure the assumptions of linear regression were satisfied, I performed a residual analysis to examine if the residuals were normally distributed with a constant variance (Faraway, 2004). The normality assumption was checked with the Kolmogorov-Smirnov (KS) test with the null hypotheses of normal distribution. The test statistic ( $D=0.32331$  and  $p\text{-value}<0.01$ ) and the visualization of the residuals in Figure 5.3 suggest a non-normal distribution. The constant variance assumption was checked with the Non-Constant Variance (NCV) test with the null hypothesis of constant variance. The NCV test suggests that the variance of the residuals is not constant (Chi-square=328011.1 and  $p\text{-value}<0.01$ ) and that heteroscedasticity was present. As shown in Table 5.3, the model has low explanatory power with an  $R$ -Squared and Adjusted  $R$ -Squared of 0.309. The model accounted for only 30.9% of the variation of the engagement scores.

**Table 5.3** Regression model with engagement scores as the dependent variable, richness scores as dependent variables, and controlled for the effects of numbers of followers, celebrity type and platform. The model has low R-Squared and violates the assumptions of linear regression.

<i>EngagementScore</i> =	<b>Est. Coef.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	0.323	0.008	41.794	<0.01*
Identity	0.056	0.005	12.005	<0.01*
Interaction	0.065	0.005	14.233	<0.01*
Visibility	-0.035	0.005	-7.534	<0.01*
Followers	0.433	0.002	181.365	<0.01*
CelebrityType	0.250	0.005	47.444	<0.01*
Platform	-0.453	0.007	-68.723	<0.01*
Residual standard error: 0.8342 on 132816 degrees of freedom				
Multiple <i>R</i> -squared: 0.3041, Adjusted <i>R</i> -squared: 0.3041				
F-statistic: 9675 on 6 and 132816 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.3** Normal plot of the residuals.

Given that the residuals were not normally distributed, I used the Box-Cox power transformation to identify an appropriate exponent ( $\lambda$ ) for transforming the dependent variable (Box & Cox, 1964). The  $\lambda$  of approximately 0.0 suggests a logarithm transformation. Thus, I constructed another model (Table 5.4) with the logarithm transformed dependent variable while the other variables remained the same.

*EngagementScore*      The logarithm transformed continuous dependent variable.

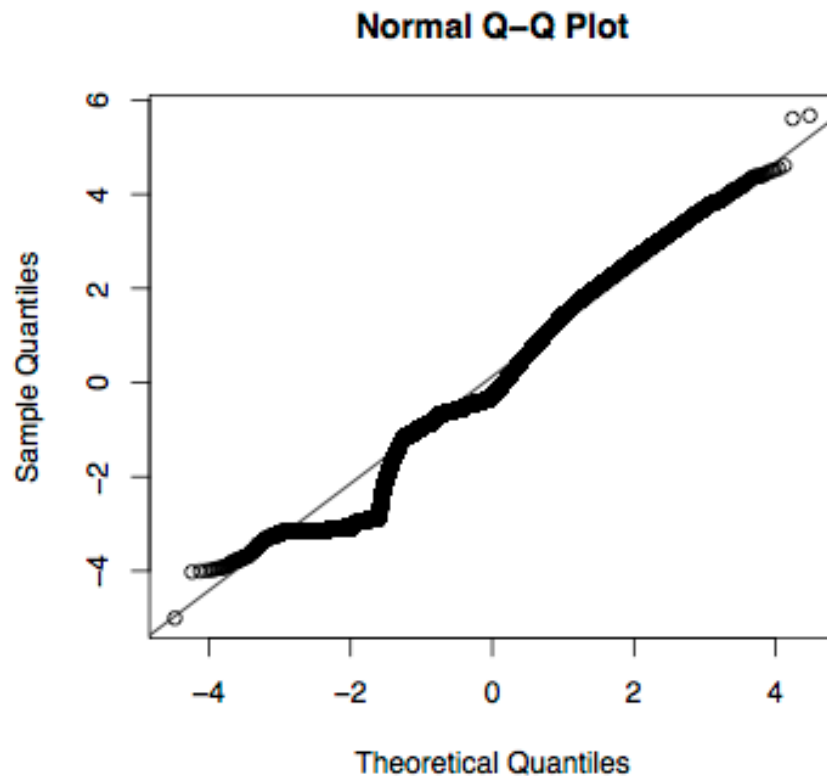
<i>Identity</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in identity measure otherwise zero.
<i>Interaction</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in interaction measure otherwise zero.
<i>Visibility</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in visibility measure otherwise zero.
<i>Followers</i>	A normalized continuous variable represents the number of followers when a post was created.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

Although the *R*-Squared value improved (*R*-Squared=0.6949), the transformation did not rectify the violation of the normality assumption as indicated by the normality test (*D*=0.13067 and *p*-value<0.01) and the visualization in Figure 5.4. The NCV test suggests that the variance of the residuals is not constant (Chi-square=4972.901 and *p*-value<0.01) and that heteroscedasticity is still present. Moreover, some VIF coefficients are higher than two, suggesting that the model suffers from multicollinearity (Faraway, 2004; James, Witten, Hastie, & Tibshirani, 2013). The high VIF coefficients are from *Followers* and *Platform*, suggesting these variables are correlated. The correlation coefficient between these two variables is 0.41.

**Table 5.4** Regression model with logarithm transformed engagement scores as the dependent variable, richness scores as dependent variables and controlled for the effects of numbers of followers, celebrity type and platform. The model has moderate R-Squared value but violates the assumptions of linear regression.

<i>EngagementScore</i> =	Est. Coef.	SE.	<i>t</i> -value	<i>p</i> -value	Inv. Log.
(Intercept)	5.269	0.012	432.244	<0.01*	194.157
Identity	0.181	0.007	24.536	<0.01*	1.199
Interaction	0.080	0.007	11.046	<0.01*	1.080
Visibility	-0.065	0.007	-8.893	<0.01*	0.937
Followers	0.881	0.004	233.519	<0.01*	2.413
CelebrityType	-2.528	0.008	-303.491	<0.01*	0.080
Platform	-3.964	0.010	-381.000	<0.01*	0.019
Residual standard error: 1.318 on 132816 degrees of freedom					
Multiple <i>R</i> -squared: 0.6949, Adjusted <i>R</i> -squared: 0.6949					
F-statistic: 5.043e+04 on 6 and 132816 <i>DF</i> , <i>p</i> -value: < 2.2e-16					





**Figure 5.4** Normal plot of the residuals.

To resolve the problem of multicollinearity, I moved the modeling in a different direction by removing *Followers* from being a control variable and re-calculating the engagement scores based on the Facebook's (2017) official engagement formula (Equation 4.2 in the main document). For Instagram, the engagement scores are re-calculated as the summation of numbers of *likes* and *comments*, divided by *number of followers* (Equation 4.3 in the main document). The calculation is similar for Twitter but the numerator is the summation of numbers of *favorites*, *replies* and *retweets*, and divided by *number of followers* (Equation 4.4 in the main document).

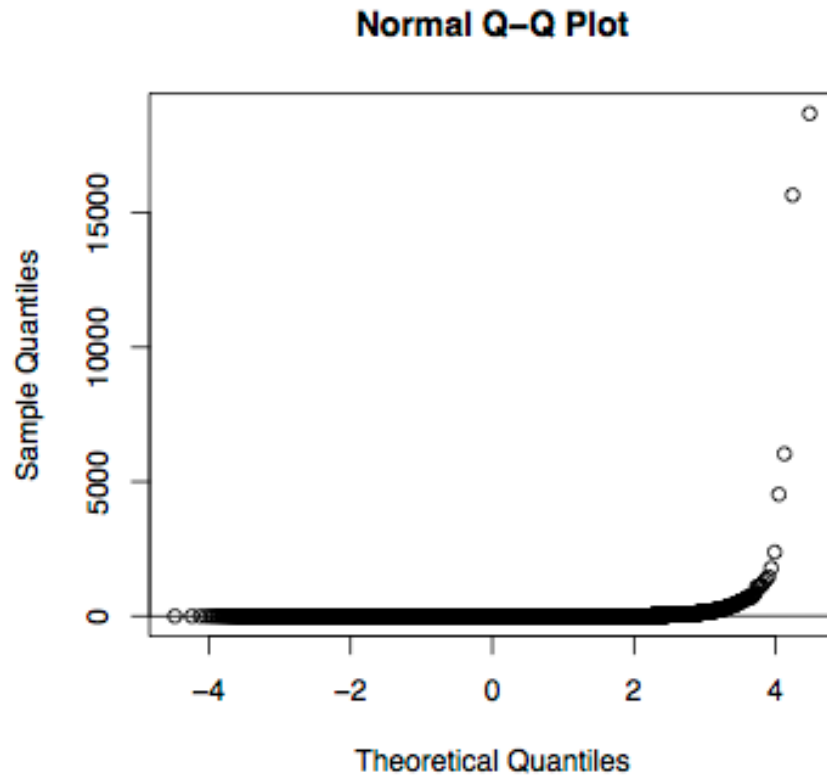
I began the modeling by using the raw updated engagement scores as the dependent variable (i.e., no transformation). The model controlled for the effects of platform and celebrity types. As before, all continuous variables were normalized to make the interpretations easier.

<i>EngagementScore</i>	The normalized continuous dependent variable.
<i>Identity</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in identity measure otherwise zero.
<i>Interaction</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in interaction measure otherwise zero.
<i>Visibility</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in visibility measure otherwise zero.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

None of the VIF coefficients is greater than two, suggesting that the model does not suffer from multicollinearity. The residuals are not normally distributed as indicated by the normality test ( $D=0.35634$  and  $p\text{-value}<0.01$ ) and the visualization in Figure 5.5. The residuals, however, have constant variance as suggested by the non-constant variance score test (Chi-square= $0.1923$  and  $p=0.661$ ). The model has an  $R$ -Squared of  $0.11017$  and an Adjusted  $R$ -Squared of  $0.19351$ .

**Table 5.5** Regression model with the updated engagement scores as the dependent variable, richness scores as dependent variables and controlled for the effects of numbers of followers, celebrity type and platform. The model has low *R*-Squared and violates some assumptions of linear regression.

<i>EngagementScore</i> =	<b>Est. Coef.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	2.707	0.377	7.177	<0.01*
Identity	0.430	0.267	1.611	0.107
Interaction	0.145	0.271	0.537	0.591
Visibility	-0.022	0.267	-0.081	0.936
CelebrityType	-2.037	0.275	-7.394	<0.01*
Platform	-0.394	0.293	-1.346	0.178
Residual standard error: 32.79 on 132817 degrees of freedom				
Multiple <i>R</i> -squared: 0.11017, Adjusted <i>R</i> -squared: 0.19351				
F-statistic: 12.05 on 5 and 132817 <i>DF</i> , <i>p</i> -value: 4.324e-12				



**Figure 5.5** Normal plot of the residuals.

As before, I used the Box-Cox power transformation to identify an appropriate  $\lambda$  for transforming the dependent variable. The lambda of 0.1 suggests a logarithm function. As such, the last version of modeling transformed the dependent variable with a logarithm function while other variables remained the same.

*EngagementScore*      The logarithm transformed continuous dependent variable.

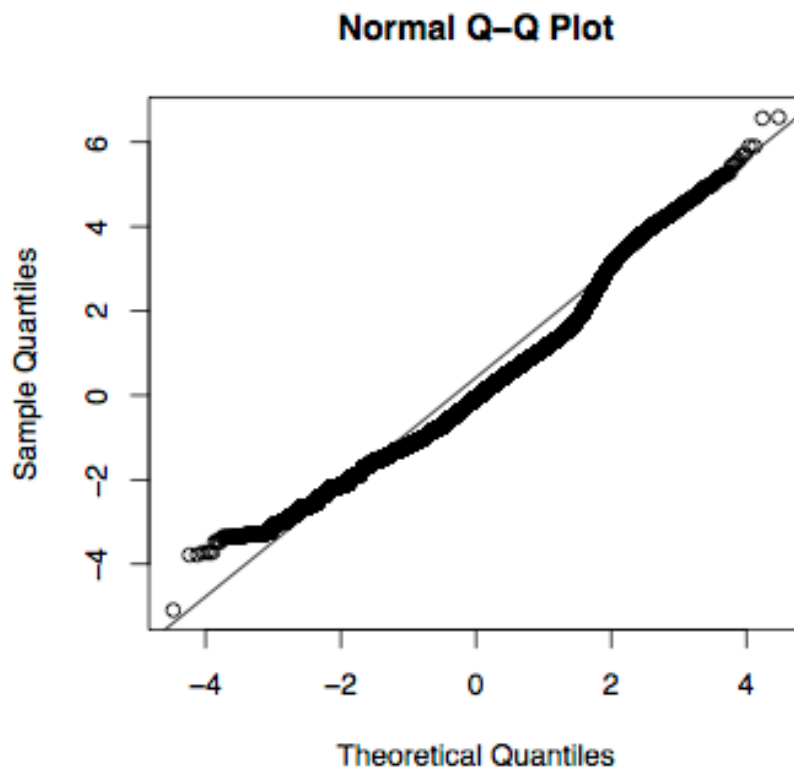
*Identity*                A Boolean variable, one if a post was categorized as *rich* in identity measure otherwise zero.

<i>Interaction</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in interaction measure otherwise zero.
<i>Visibility</i>	A Boolean variable, one if a post was categorized as <i>rich</i> in visibility measure otherwise zero.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

The VIF coefficients are between 1.01 to 1.06, suggesting that none of the independent variables are strongly correlated and that the model does not suffer from multicollinearity. The normality test statistic ( $D=0.0024$  and  $p\text{-value}=0.3214$ ) suggests that the residuals are normally distributed. The Non-Constant Variance test suggests that the residuals are homoscedastic i.e., they have constant variance ( $\text{Chi-square}=0.011$ ,  $p\text{-value}=0.9157$ ). Both  $R\text{-Squared}$  and Adjusted  $R\text{-Squared}$  are 0.8026 indicating the model includes only relevant predictors and could explain the variation well. All predictors are statistically significant with  $p\text{-value}$  of less than 0.01.

**Table 5.6** Regression model with logarithm transformed engagement scores as the dependent variable, richness scores as dependent variables and controlled for the effects of celebrity type and platform. The model has high *R*-Squared and satisfies the assumptions of linear regression.

	<b>Est. Coef.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>	<b>Inv. Log.</b>
(Intercept)	0.738	0.010	70.433	<0.01*	2.092
Identity	0.076	0.008	9.873	<0.01*	1.079
Interaction	0.107	0.008	14.282	<0.01*	1.113
Visibility	-0.054	0.007	-7.210	<0.01*	0.948
CelebrityType	-1.456	0.008	-190.286	<0.01*	0.233
Platform	-3.366	0.008	-413.548	<0.01*	0.035
Residual standard error: 0.9108 on 132817 degrees of freedom					
Multiple <i>R</i> -squared: 0.8026, Adjusted <i>R</i> -squared: 0.8026					
<i>F</i> -statistic: 4.97e+04 on 5 and 132817 <i>DF</i> , <i>p</i> -value: < 2.2e-16					



**Figure 5.6** Normal plot of the residuals.

Table 5.7 summarizes the performance of the four models and shows that the last model outperforms all others by satisfying the assumptions of linear regression but also has the highest explanatory power. As such, the last model was chosen. Given that the dependent variable was transformed with the logarithm function, the coefficients of the model are presented as an inverse logarithm. The inverse log coefficient of identity richness scores of 1.08 suggests that one unit increase in the identity richness score increases the engagement score by 8%. For interaction richness, the inverse log coefficient of 1.11 suggests that one unit increase in interaction richness increases the engagement score by 11%. These suggest that the audience tends to engage when

celebrities open access to their life or show an attempt to interact with the public. On the other hand, the visibility's coefficient of 0.95 suggests that one unit increase in visibility score reduces the engagement score by 5% (a unit increase in score results in 95% engagement score thus reduces by  $100-95=5\%$ ). This is an interesting finding which suggests that the more celebrities try to promote themselves, the less the audience tends to engage. For the celebrity type with mainstream celebrity as a base state, the coefficient of 0.23 suggests that engagement scores are higher for Internet celebrities. The coefficient of the platform variable is 0.035 suggests that the engagement score is higher on Instagram.

I answer the third research question concerning how the audiences respond to different microcelebrity strategies that, when using engagement scores as a measure of audience response, the audiences tend to be more engaged with identity- and interaction-rich posts but less engaged with visibility-rich posts.



**Table 5.7** Summary of models' performance. This shows that the last model outperforms all others with the highest  $R$ -Squared and Adjusted  $R$ -Squared while satisfying the assumptions of linear regression.

	<b>Model 1.1</b>	<b>Model 1.2</b>	<b>Model 1.3</b>	<b>Model 1.4</b>
Residuals are normally distributed.	No	No	No	Yes
Residuals are homoscedastic.	No	No	No	Yes
No multicollinearity.	No	No	Yes	Yes
$R$ -Squared, Adjusted $R$ -Squared.	0.3041, 0.3041	0.6949, 0.6949	0.1101, 0.1935	0.8026, 0.8026

***Model 2: Mean engagement scores and variance of richness scores***

To examine the relationships between engagement scores and *changes* in microcelebrity strategies, I operationalized the *changes* by calculating the *variance* of richness scores for each user. Each observation represents a user whose engagement scores of all posts on each platform were aggregated using the *mean*. For each richness dimension, I calculated the *variance* of the richness scores of all posts on each platform for each user. Each observation consists of the mean engagement score, three variance scores, the mean number of followers, the platform where the aggregated posts occur and celebrity type. Note that each of the 78 users appears in the aggregated dataset twice (one for each platform). This model aims to explore if users with more (or less) variance in their richness scores tend to have more (or less) mean engagement. Thus, the

regression explains how the aggregated variance in celebrities' scores is related to their mean engagement scores. The header for the dataset of 156 records is:

Mean engagement score	Identity variance score	Interaction variance score	Visibility variance score	Mean number of followers	Platform	Celebrity type
-----------------------	-------------------------	----------------------------	---------------------------	--------------------------	----------	----------------

As before, I experimented with different versions of modeling before obtaining the final model. The first version used the number of followers as a control variable, and calculated the engagement scores as the summation of the numbers of *likes* and *comments* for Instagram posts; or *likes*, *replies* and *retweets* for Twitter data. I constructed the model using the raw mean engagement scores (i.e., no transformation) as the dependent variable. The independent variables were the variances of richness scores in three dimensions. The model controlled for the effects of number of followers, platforms and celebrity types. Note that all but the dummy variables were normalized to make the interpretations easier.

*EngagementScore*      The normalized continuous dependent variable represents the mean values of the engagement scores of each user.

*IdentityVar*            A continuous variable represents the variance of identity scores of each user.

*InteractionVar*        A continuous variable represents the variance of interaction scores of each user.

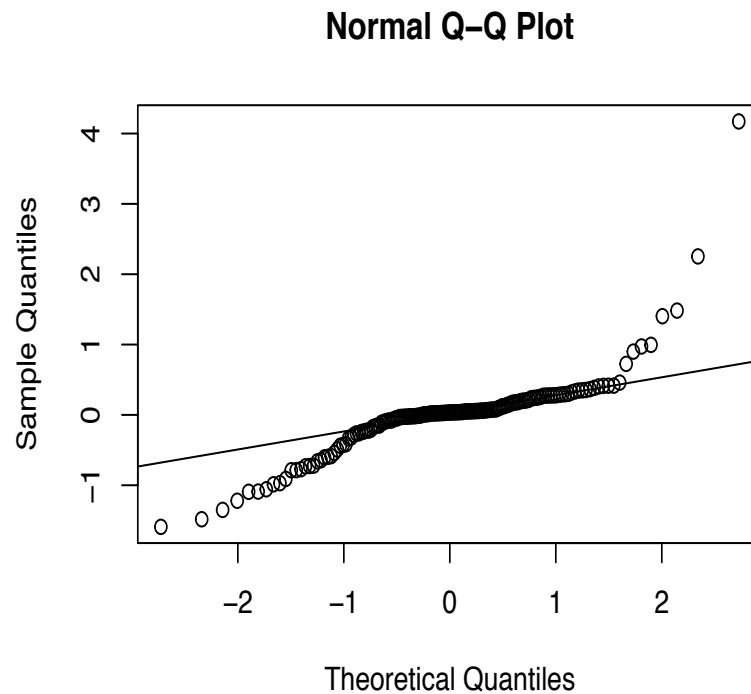
*VisibilityVar*         A continuous variable represents the variance of visibility scores of each user.

<i>FollowerMean</i>	A normalized continuous variable represents the mean number of followers of a user.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

I performed a residual analysis to examine if the residuals were normally distributed with a constant variance (Faraway, 2004). The normality test statistic  $D$  of 0.281 and  $p$ -value $<0.01$  along with the visualization of the residuals in Figure 5.7 suggest a non-normal distribution. The NCV test suggests that the variance of the residuals is not constant (Chi-square=390.106 and  $p<0.01$ ) and that heteroscedasticity is present. The model has moderate explanatory power with an  $R$ -Squared of 0.6458 and an Adjusted  $R$ -Squared of 0.6316.

**Table 5.8** Regression model with mean engagement scores as the dependent variable, variances of richness scores as dependent variables, and controlled for the effects of mean followers, celebrity type and platform. The model has moderate *R*-Squared and violates the assumptions of linear regression.

<b>EngagementScore =</b>	<b>Est. Coeff.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	1.153	0.277	4.157	<0.01*
IdentityVar	-0.073	1.019	-0.072	0.943
InteractionVar	-3.475	1.261	-2.755	<0.01*
VisibilityVar	-2.416	0.947	-2.551	0.012*
FollowerMean	0.607	0.049	12.279	<0.01*
CelebrityType	0.680	0.099	6.894	<0.01*
Platform	-0.409	0.122	-3.340	<0.01*
Residual standard error: 0.607 on 149 degrees of freedom				
Multiple <i>R</i> -squared: 0.6458, Adjusted <i>R</i> -squared: 0.6316				
F-statistic: 45.28 on 6 and 149 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.7** Normal plot of the residuals.

Given that the residuals are not normally distributed, I used the Box-Cox power transformation to obtain an appropriate  $\lambda$ . The  $\lambda$  of approximately 0.07 suggests a logarithm function. Thus, I constructed another model (Table 5.9) with the logarithm transformed dependent variable while other variables remained the same.

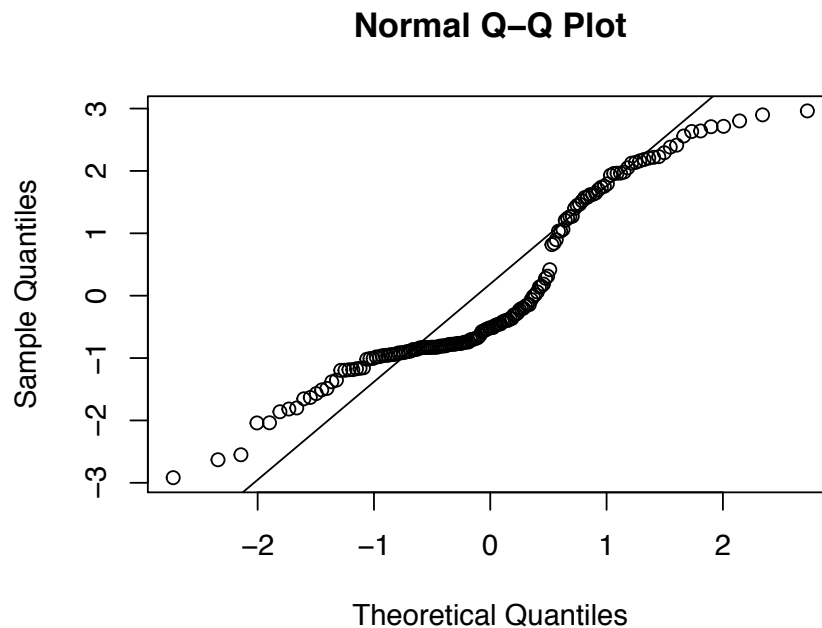
*EngagementScore*      The logarithm transformed dependent variable represents the mean engagement scores of each user.

<i>IdentityVar</i>	A continuous variable represents the variance of identity scores of each user.
<i>InteractionVar</i>	A continuous variable represents the variance of interaction scores of each user.
<i>VisibilityVar</i>	A continuous variable represents the variance of visibility scores of each user.
<i>FollowerMean</i>	A normalized continuous variable represents the mean number of followers of a user.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

The transformation, however, did not rectify the violation of the normality assumption as indicated by the normality test ( $D=0.219$  and  $p\text{-value}<0.01$ ) and the visualization in Figure 5.8. The NCV test suggests that the variance of the residuals is not constant (Chi-square=4.450 and  $p\text{-value}=0.03$ ) and that heteroscedasticity is still present. None of the VIF coefficients are higher than two, suggest that the model does not suffers from multicollinearity. The model has high explanatory power with an  $R\text{-Squared}$  of 0.80 and an Adjusted  $R\text{-Squared}$  of 0.792.

**Table 5.9** Regression model with logarithm transformed mean engagement scores as the dependent variable, variances of richness scores as dependent variables, and controlled for the effects of mean followers, celebrity type and platform. The model has high *R*-Squared but violates some assumptions of linear regression.

<b>EngagementScore =</b>	<b>Est. Coeff.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>	<b>Inv. Log.</b>
(Intercept)	6.349	0.633	10.033	<0.01*	571.748
IdentityVar	2.953	2.326	1.269	0.206	19.169
InteractionVar	-7.015	2.879	-2.437	0.016*	0.001
VisibilityVar	0.828	2.162	0.383	0.702	0.437
FollowerMean	1.458	0.113	12.927	<0.01*	4.299
CelebrityType	-2.196	0.225	-9.750	<0.01*	0.111
Platform	-3.243	0.279	-11.605	<0.01*	0.039
Residual standard error: 1.385 on 149 degrees of freedom					
Multiple R-squared: 0.800, Adjusted R-squared: 0.792					
F-statistic: 99.36 on 6 and 149 <i>DF</i> , <i>p</i> -value: < 2.2e-16					



**Figure 5.8** Normal plot of the residuals.

Another version of modelling removed *FollowerMean* from being a control variable and re-calculated the engagement score using Equation 4.3 and 4.4 (in the main document). Recall that the engagement scores for Instagram are the summation of numbers of *likes* and *comments*, divided by *number of followers* (Equation 4.3 in the main document). The calculation is similar for Twitter but the numerator is the summation of numbers of *favorites*, *replies* and *retweets*, and divided by *number of followers* (Equation 4.4 in the main document).

I began the modeling by using the raw updated mean engagement scores as the dependent variable (i.e., no transformation). The model controlled for the effects of platforms and celebrity types. As before, all continuous variables were normalized to make the interpretations easier.

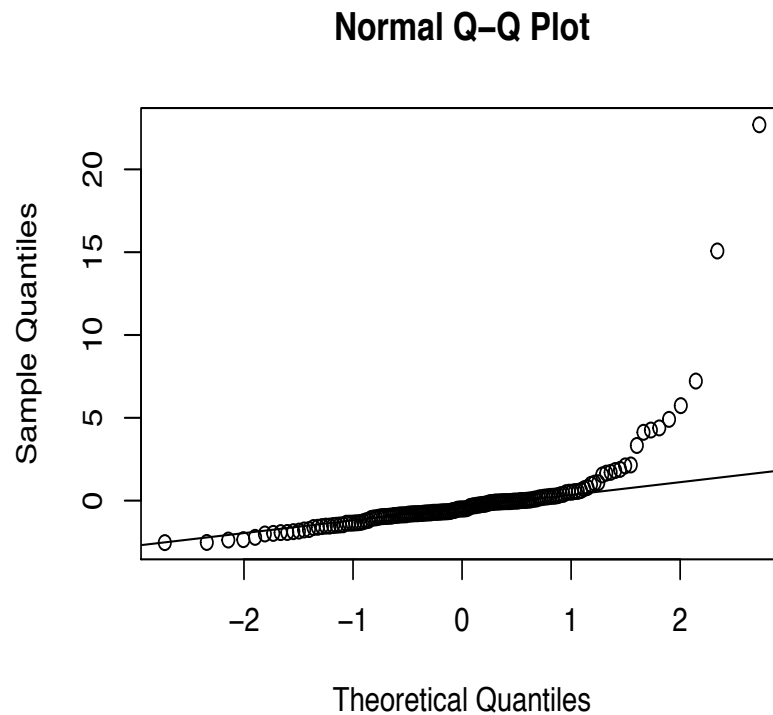


<i>EngagementScore</i>	The normalized continuous dependent variable represents the mean values of the engagement scores of each user.
<i>IdentityVar</i>	A continuous variable represents the variance of identity scores of each user.
<i>InteractionVar</i>	A continuous variable represents the variance of interaction scores of each user.
<i>VisibilityVar</i>	A continuous variable represents the variance of visibility scores of each user.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

None of the VIF coefficients is greater than two, suggesting that the model does not suffer from multicollinearity. The residuals have non-constant variance (Chi-square=7.806 and  $p$ -value=0.005) and not normally distributed as indicated by the normality test ( $D$ =0.223 and  $p$ -value<0.01) and the visualization in Figure 5.9. The model has low explanatory power with an  $R$ -Squared of 0.187 and an Adjusted  $R$ -Squared of 0.160.

**Table 5.10** Regression model with updated mean engagement scores as the dependent variable, variances of richness scores as dependent variables, and controlled for the effects of celebrity type and platform. The model has low R-Squared and violates the assumptions of linear regression.

<b>EngagementScore =</b>	<b>Est. Coeff.</b>	<b>SE.</b>	<b>t-value</b>	<b>p-value</b>
(Intercept)	4.286	1.219	3.517	<0.01*
IdentityVar	5.359	4.467	1.199	0.232
InteractionVar	-10.863	5.484	-1.981	0.049*
VisibilityVar	-2.123	4.174	-0.509	0.612
CelebrityType	-0.675	0.435	-1.553	0.123
Platform	-1.651	0.538	-3.067	<0.01*
Residual standard error: 2.675 on 150 degrees of freedom Multiple R-squared: 0.1872, Adjusted R-squared: 0.1601 F-statistic: 6.908 on 5 and 150 DF, p-value: 0.000007854				



**Figure 5.9** Normal plots of the residuals.

As before, I used the Box-Cox power transformation which suggested a lambda of 0.1 or a logarithm transformation. The last version of modeling thus transformed the dependent variable with a logarithm function while other variables remained the same.

*EngagementScore*      The logarithm transformed continuous dependent variable represents the mean values of the engagement scores of each user.

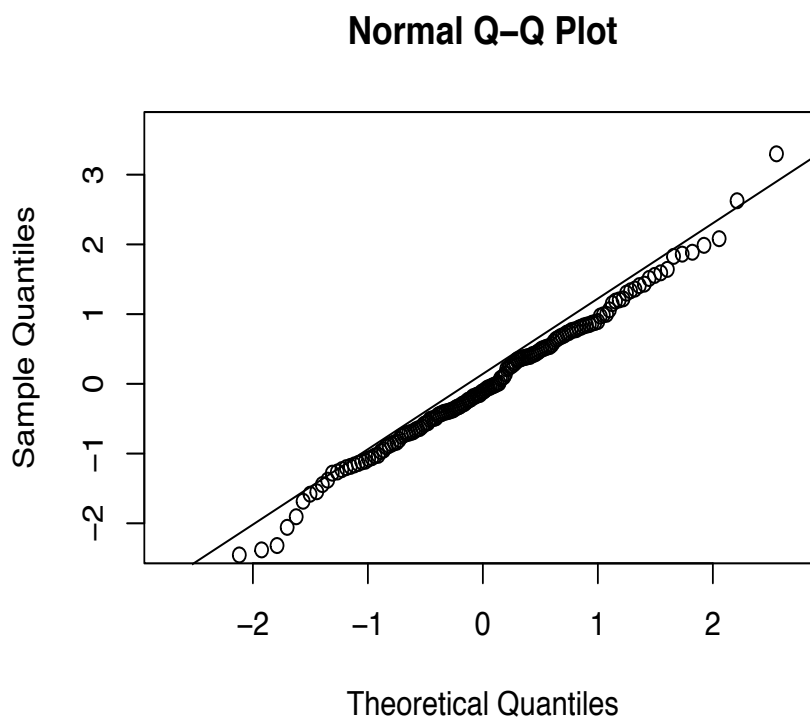
*IdentityVar*            A continuous variable represents the variance of identity scores of each user.

<i>InteractionVar</i>	A continuous variable represents the variance of interaction scores of each user.
<i>VisibilityVar</i>	A continuous variable represents the variance of visibility scores of each user.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state; one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state; one if a post was a tweet otherwise zero.

The model is presented in Table 5.11 with an *R*-Squared of 0.8354 and Adjusted *R*-Squared of 0.8299. This indicates that the model explains the variation in response very well. The normality test ( $D=0.0776$ ,  $p$ -value=0.3034) and Non-Constant Variance test (Chi-square=0.5000 and  $p$ -value=0.4795) along with the visualization (Figure 5.10) show that the residuals are normally distributed with constant variance. The VIF coefficients are between 1.03 to 1.36, indicating that the model does not suffer from multicollinearity.

**Table 5.11** Regression model with logarithm transformed mean engagement scores as the dependent variable, variances of richness scores as dependent variables, and controlled for the effects of celebrity type and platform. The model has high *R*-Squared and satisfies the assumptions of linear regression.

<b>EngagementScore =</b>	<b>Est. Coeff.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>	<b>Inv. Log.</b>
(Intercept)	1.252	0.356	3.522	<0.01*	3.498
IdentityVar	0.434	2.194	0.198	0.844	1.543
InteractionVar	-2.672	1.303	-2.050	0.042*	0.069
VisibilityVar	-1.047	1.218	-0.860	0.391	0.351
CelebrityType	-1.135	0.127	-8.945	<0.01*	0.321
Platform	-2.861	0.157	-18.215	<0.01*	0.057
Residual standard error: 0.7804 on 150 degrees of freedom					
Multiple <i>R</i> -squared: 0.8354, Adjusted <i>R</i> -squared: 0.8299					
<i>F</i> -statistic: 152.2 on 5 and 150 <i>DF</i> , <i>p</i> -value: < 2.2e-16					



**Figure 5.10** Normal plot of the residuals.

The four models are summarized in Table 5.12, which shows that the last model outperforms all others by satisfying the assumptions of linear regression but also has the highest explanatory power. As such, this model was chosen. The results indicate that among the richness score, only the interaction variance score has a significant negative effect with the  $p$ -value of 0.042. Celebrity type and platform variables are also statistically significant. Given that the dependent variable was transformed with the logarithm function, the coefficients of the model are presented as inverse logarithm. The inverse log coefficient of the variance in interaction scores of 0.069 suggests that one unit increase in variance of the interaction scores decreases the engagement score by 93.6% (a unit increase in score variance results in 6.9% engagement score

thus reduced by  $100 - 6.9 = 93.6\%$ ). This suggests that the audience tends to be less engaged with a celebrity whose interactional level is not consistent. For the celebrity type with mainstream celebrity as a base state, the coefficient of 0.321 suggests that engagement scores are higher for Internet celebrities. The coefficient of the platform variable is 0.057 suggests that the engagement scores are higher on Instagram.

I answer the fourth research question concerning how the audiences respond to the changes in microcelebrity strategies that only the consistency in interaction richness has effects on the audience's engagement.

**Table 5.12** Summary of models' performance. This shows that the last model outperforms all others with the highest *R*-Squared and Adjusted *R*-Squared while satisfying the assumptions of linear regression.

	<b>Model 2.1</b>	<b>Model 2.2</b>	<b>Model 2.3</b>	<b>Model 2.4</b>
Residuals are normally distributed.	No	No	No	Yes
Residuals are homoscedastic.	No	No	No	Yes
No multicollinearity.	No	Yes	Yes	Yes
<i>R</i> -Squared, Adjusted <i>R</i> -Squared.	0.4723, 0.4729	0.800, 0.792	0.1872, 0.1601	0.8354, 0.8299

**Model 3: Follower changes and mean richness scores**

The other measure of an audience's responses is the changes in numbers of followers and so this analysis was performed at the user level. For each user, I calculated the change in numbers of followers as the difference between the numbers of followers recorded on the last and first day of the data collection, and the richness scores were aggregated using the *mean*. Each observation represents a user and consists of the change in numbers of followers, three aggregated richness scores, number of followers on the first day of data collection, the total number of posts, platform where the aggregated posts occur, and celebrity type. Note that each of the 78 users appears in the aggregated dataset twice (one for each platform). The header for the dataset of 156 records is:

Follower change	Mean identity score	Mean interaction score	Mean visibility score	Number of followers	Number of posts	Platform	Celebrity type
-----------------	---------------------	------------------------	-----------------------	---------------------	-----------------	----------	----------------

To examine the relationship between changes in numbers of followers and richness scores, I constructed a model using the change in numbers of followers as the dependent variable and the aggregated richness scores in three dimensions as independent variables. The model controlled for the effects of number of followers on the first day of data collection, number of posts, celebrity types with mainstream as a base state and platforms with Twitter as a base state. Note that all but the dummy variables were normalized to make the interpretations easier.

*FollowerChange*      The normalized continuous dependent variable.

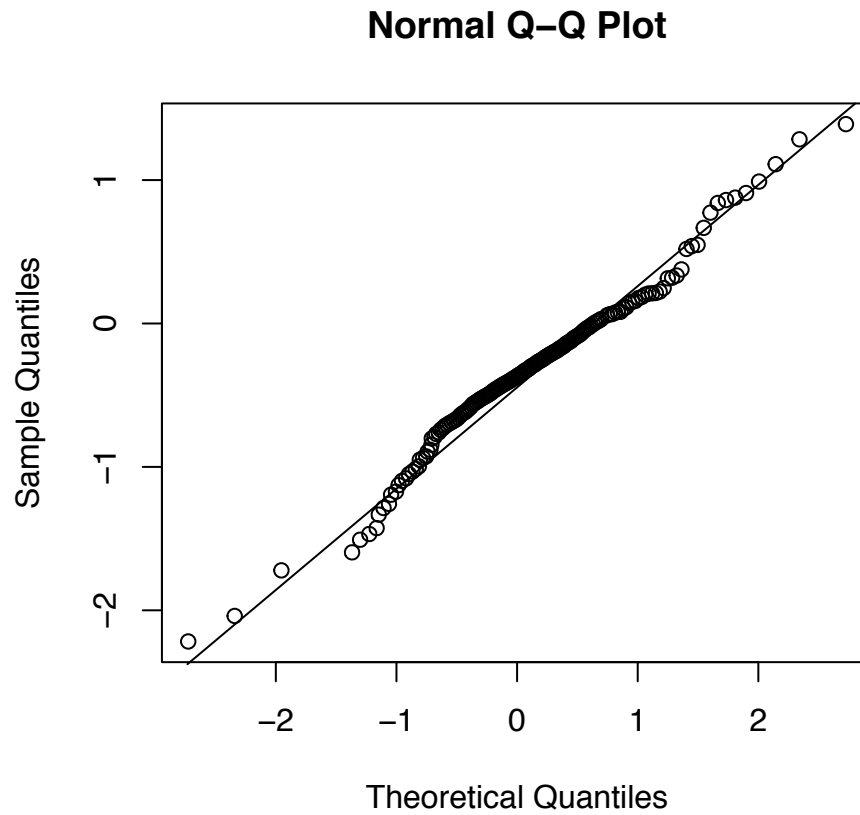


<i>IdentityMean</i>	A normalized continuous variable represents the mean of identity richness scores of all posts of a user.
<i>InteractionMean</i>	A normalized continuous variable represents the mean of interaction richness scores of all posts of a user.
<i>VisibilityMean</i>	A normalized continuous variable represents the mean of visibility richness scores of all posts of a user.
<i>FollowerStart</i>	A normalized continuous variable represents the number of followers recorded on the first day of the collection.
<i>TotalPost</i>	A normalized continuous variable represents the total number of posts.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state: one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state: one if a post was a tweet otherwise zero.

The model is presented in Table 5.13 with an *R*-Squared of 0.8416 and Adjusted *R*-Squared of 0.8341. This indicates that the model explains the variation in response very well. The normality and Non-Constant Variance tests along with the visualization (Figure 5.11) show that the residuals are normally distributed with a constant variance. The Variance Inflation Factor (VIF) coefficients are all lower than 2, indicating that the model does not suffer from multicollinearity. The results also indicate that none of the richness variables are significant with *p*-values all over 0.05. All control variables but *TotalPost* are statistically significant at a 95% confidence level.

**Table 5.13** Regression model with follower changes as the dependent variable, mean richness scores as dependent variables, and controlled for the effects of followers, number of posts, celebrity type and platform. The model has high *R*-Squared and satisfies the assumptions of linear regression.

<b>FollowerChange=</b>	<b>Est. Coef.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	-0.013	0.077	-0.174	0.862
IdentityMean	0.027	0.071	0.380	0.704
InteractionMean	-0.032	0.039	-0.825	0.411
VisibilityMean	-0.011	0.042	-0.267	0.790
FollowerStart	0.883	0.033	26.489	<0.01*
TotalPost	0.006	0.036	0.163	0.871
CelebrityType	0.339	0.071	4.751	<0.01*
Platform	-0.312	0.135	-2.306	0.023*
Residual standard error: 0.4073 on 148 degrees of freedom				
Multiple R-squared: 0.8416, Adjusted R-squared: 0.8341				
<i>F</i> -statistic: 112.3 on 7 and 148 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.11** Normal plot of the residuals.

Given that *TotalPost* is not statistically significant at a 95% confidence level, another version of the modeling removed *TotalPost* while all other variables remained the same.

*FollowerChange*      The normalized continuous dependent variable.

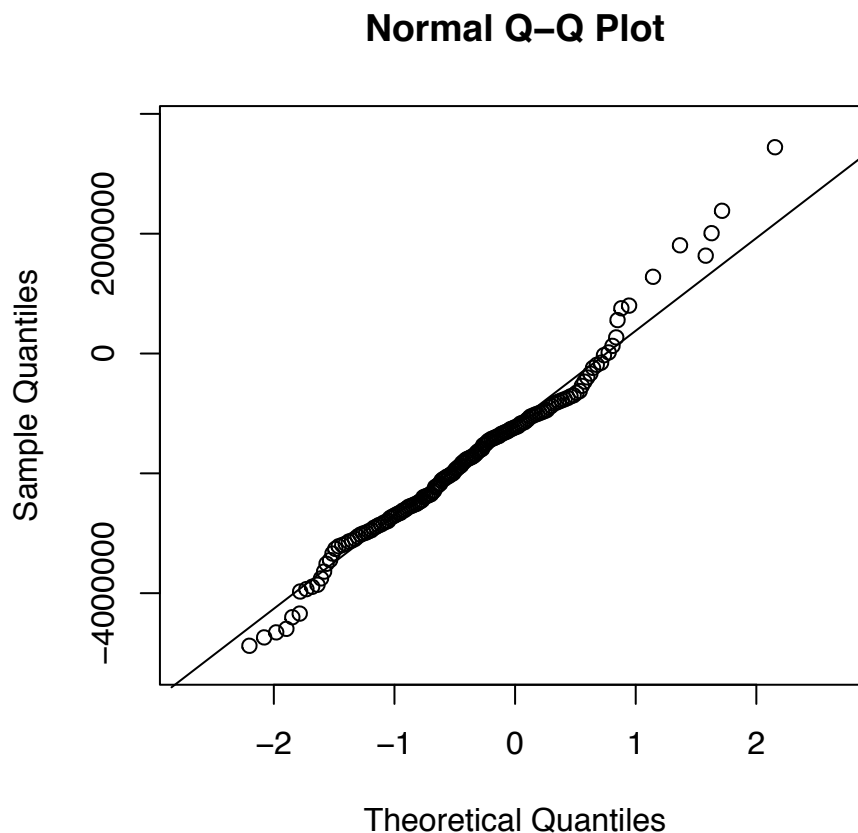
*IdentityMean*      A normalized continuous variable represents the mean of identity richness scores of all posts of a user.

<i>InteractionMean</i>	A normalized continuous variable represents the mean of interaction richness scores of all posts of a user.
<i>VisibilityMean</i>	A normalized continuous variable represents the mean of visibility richness scores of all posts of a user.
<i>FollowerStart</i>	A normalized continuous variable represents the number of followers recorded on the first day of the collection.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state: one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state: one if a post was a tweet otherwise zero.

The model is presented in Table 5.14 with an *R*-Squared of 0.8416 and Adjusted *R*-Squared of 0.8352. This indicates that the model explains the variation in response very well. The normality and Non-Constant Variance tests along with the visualization (Figure 5.12) show that the residuals are normally distributed with a constant variance. The Variance Inflation Factor (VIF) coefficients are all lower than 2, indicating that the model does not suffer from multicollinearity. The results also indicate that none of the richness variables are significant with *p*-values all over 0.05. All control variables are statistically significant at a 95% confidence interval.

**Table 5.14** Regression model with follower changes as the dependent variable, mean richness scores as dependent variables, and controlled for the effects of followers, celebrity type and platform. The model has high *R*-Squared and satisfies the assumptions of linear regression.

<b>FollowerChange=</b>	<b>Est. Coef.</b>	<b>SE.</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	-0.016	0.074	-0.220	0.826
IdentityMean	0.027	0.071	0.380	0.704
InteractionMean	-0.032	0.039	-0.821	0.413
VisibilityMean	-0.011	0.042	-0.273	0.785
FollowerStart	0.883	0.033	26.734	<0.01*
CelebrityType	0.341	0.070	4.893	<0.01*
Platform	-0.308	0.133	-2.318	0.022*
Residual standard error: 0.406 on 149 degrees of freedom				
Multiple R-squared: 0.8416, Adjusted R-squared: 0.8352				
<i>F</i> -statistic: 131.9 on 6 and 149 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.12** Normal plot of the residuals.

Table 5.15 summarizes the performance of the two models, which shows that the two models perform equally well with the similar  $R$ -Squared and slightly different Adjusted  $R$ -Squared. The last model, however, is simpler with less number of predictors. As such, this model was chosen. None of the richness variables is statistically significant at a 95% confidence interval, suggesting that we could not detect the effects of the richness scores on the changes in numbers of followers. On the other hand, all control variables are statistically significant. The positive coefficients of *FollowerStart* and *CelebrityType* suggests that the more followers a user

starts with, the more followers the user is likely to get, and that mainstream celebrities are more likely to gain more followers than Internet celebrities. Lastly, the negative coefficient of *Platform* suggests that follower changes are generally higher on Instagram.

In summary, using changes in followers as a measure of audience response, I answer the third research question that microcelebrity strategies have no effects on the changes in the number of followers.

**Table 5.15** Summary of models' performance. This shows that the last model is preferable because it is simpler but has comparable performance.

	<b>Model 3.1</b>	<b>Model 3.2</b>
Residuals are normally distributed.	Yes	Yes
Residuals are homoscedastic.	Yes	Yes
No multicollinearity.	Yes	Yes
<i>R</i> -Squared, Adjusted <i>R</i> -Squared.	0.8416, 0.8341	0.8416, 0.8352

***Model 4: Follower changes and variances in richness scores***

This analysis was performed at the user level to examine the relationships between changes in followers and *changes* in microcelebrity strategies. As before, I operationalized the *changes in strategies* by calculating the *variance* of richness scores for each user. Each observation represents a user whose changes in numbers of followers were calculated as the difference between the numbers of followers recorded on the last and first day of data collection. For each richness dimension, I calculated the *variance* of the richness scores of all posts on each platform for each user. Each observation consists of the change in numbers of followers, three variance scores, the mean number of followers, the total number of posts, the platform where the aggregated posts occur and celebrity type. Note that each user appears in the aggregated dataset twice (one for each platform). This model aims to explore if users with more (or less) variance in their richness scores tend to have more (or less) follower changes. Thus, the regression explains how the aggregated variance in celebrities' scores is related to their changes in numbers of followers. The header for the dataset of 156 records is:

Follower change	Identity variance score	Interaction variance score	Visibility variance score	Number of followers	Number of posts	Platform	Celebrity type
-----------------	-------------------------	----------------------------	---------------------------	---------------------	-----------------	----------	----------------

The dependent variable is the changes in followers and the predictors are the variances of richness scores in three dimensions. As before, I controlled for the effects of number of followers on the first day of data collection, number of posts, celebrity types using mainstream as a base state and platforms using Twitter as a base state. All but the dummy variables were normalized to make the interpretations easier. The variables are as follow.



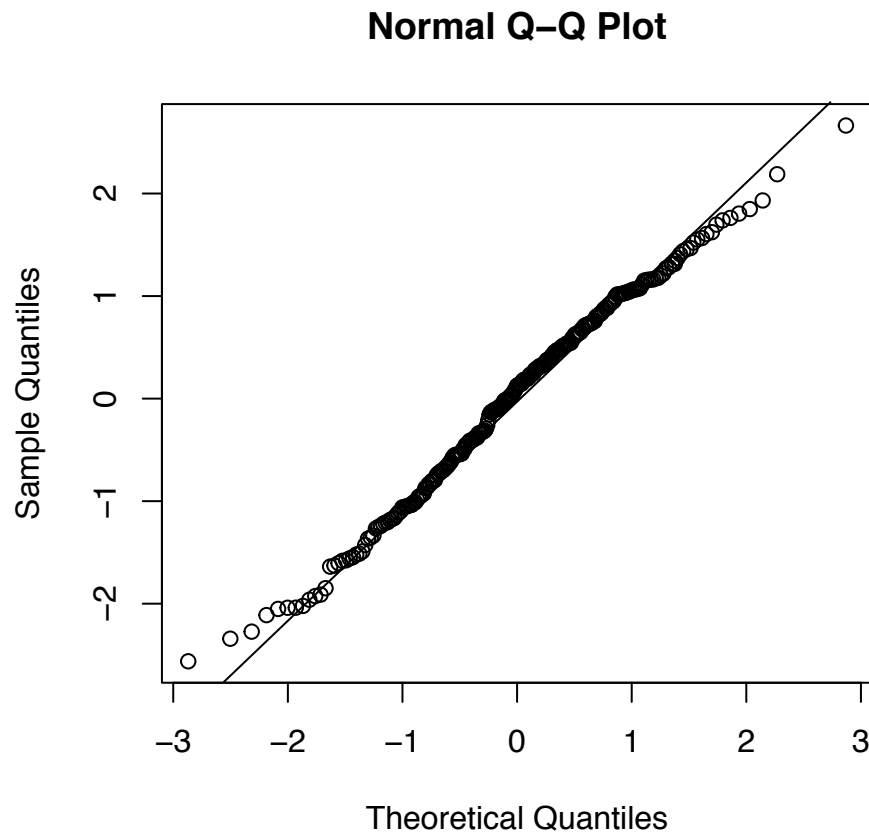
<i>FollowerChange</i>	The normalized continuous dependent variable.
<i>IdentityVar</i>	A normalized continuous variable represents the variance of identity scores of each user.
<i>InteractionVar</i>	A normalized continuous variable represents the variance of interaction scores of each user.
<i>VisibilityVar</i>	A normalized continuous variable represents the variance of visibility scores of each user.
<i>FollowerStart</i>	A normalized continuous variable represents the number of followers recorded on the first day of the collection.
<i>TotalPost</i>	A normalized continuous variable represents the total number of posts.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state: one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state: one if a post was a tweet otherwise zero.

The model is presented in Table 5.16 with an *R*-Squared of 0.8428 and Adjusted *R*-Squared of 0.8354. This indicates that the model explains the variation in response very well. The normality and Non-Constant Variance tests along with the visualization (Figure 5.13) show that the residuals are normally distributed with a constant variance. The Variance Inflation Factor (VIF) coefficients are all lower than 2, indicating that the model does not suffer from multicollinearity. The results also indicate that amongst the richness scores, all predictors except

the variance score of identity richness are significant ( $p$ -value  $< 0.05$ ). All control variables but *TotalPost* are also statistically significant.

**Table 5.16** Regression model with follower changes as the dependent variable, variances of richness scores as dependent variables, and controlled for the effects of followers, number of posts, celebrity type and platform. The model has high  $R$ -Squared and satisfies the assumptions of linear regression.

<b>FollowerChange=</b>	<b>Est. Coef.</b>	<b>SE</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	-0.001	0.062	-0.012	0.990
IdentityVar	-0.043	0.047	-0.903	0.368
InteractionVar	-0.220	0.041	-5.360	<0.01*
VisibilityVar	0.284	0.038	7.530	<0.01*
FollowerStart	0.884	0.033	26.566	<0.01*
TotalPost	0.003	0.036	0.085	0.932
CelebrityType	0.350	0.067	5.210	<0.01*
Platform	-0.348	0.085	-4.091	<0.01*
Residual standard error: 0.4057 on 148 degrees of freedom				
Multiple R-squared: 0.8428, Adjusted R-squared: 0.8354				
<i>F</i> -statistic: 113.4 on 7 and 148 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.13** Normal plot of the residuals.

Given that *TotalPost* is not statistically significant at a 95% confidence level, I constructed another model by removing *TotalPost* while all other variables remained the same. The variables are as follow.

*FollowerChange*      The normalized continuous dependent variable.

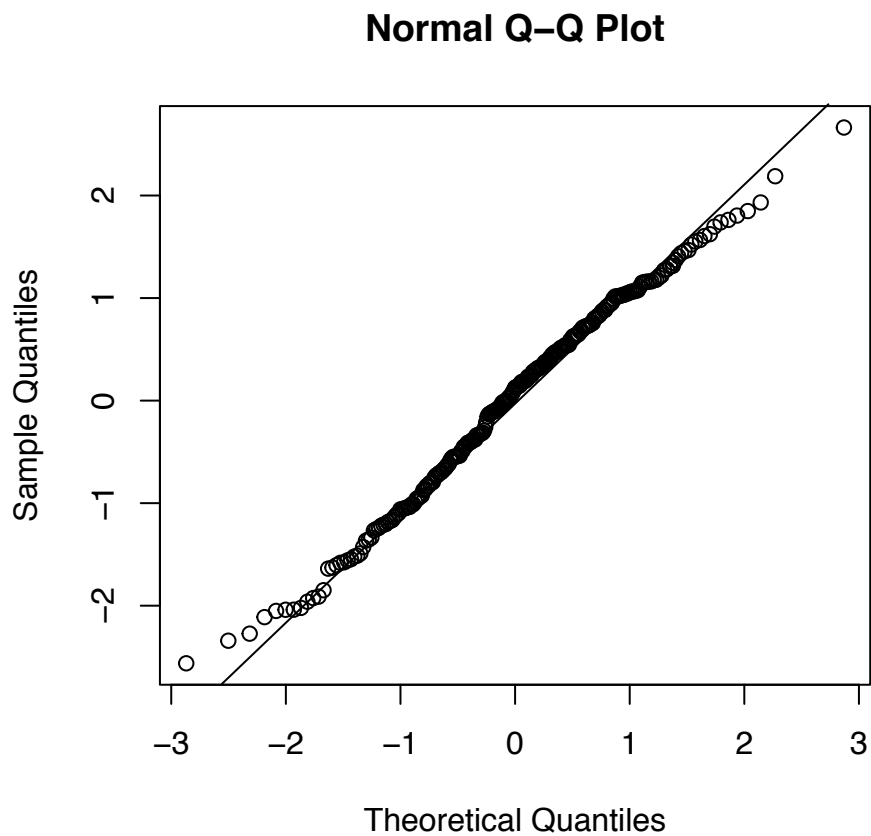
*IdentityVar*            A normalized continuous variable represents the variance of identity scores of each user.

<i>InteractionVar</i>	A normalized continuous variable represents the variance of interaction scores of each user.
<i>VisibilityVar</i>	A normalized continuous variable represents the variance of visibility scores of each user.
<i>FollowerStart</i>	A normalized continuous variable represents the number of followers recorded on the first day of the collection.
<i>CelebrityType</i>	A Boolean variable with mainstream as a base state: one if an account was a mainstream celebrity otherwise zero.
<i>Platform</i>	A Boolean variable with Twitter as a base state: one if a post was a tweet otherwise zero.

The model is presented in Table 5.17 with an *R*-Squared of 0.8428 and Adjusted *R*-Squared of 0.8365. This indicates that the model explains the variation in response very well. The normality and Non-Constant Variance tests along with the visualization (Figure 5.14) show that the residuals are normally distributed with a constant variance. The Variance Inflation Factor (VIF) coefficients are all lower than 2, indicating that the model does not suffer from multicollinearity. The results also indicate that amongst the richness scores, all predictors except the variance score of identity richness are significant ( $p$ -value < 0.05). All control variables are also statistically significant.

**Table 5.17** Regression model with follower changes as the dependent variable, variances of richness scores as independent variables, and controlled for the effects of followers, celebrity type and platform. The model has high *R*-Squared and satisfies the assumptions of linear regression.

<b>FollowerChange=</b>	<b>Est. Coef.</b>	<b>SE</b>	<b><i>t</i>-value</b>	<b><i>p</i>-value</b>
(Intercept)	-0.002	0.059	-0.039	0.969
IdentityVar	-0.043	0.047	-0.904	0.367
InteractionVar	-0.220	0.041	-5.360	<0.01*
VisibilityVar	0.284	0.038	7.530	<0.01*
FollowerStart	0.883	0.033	26.823	<0.01*
CelebrityType	0.351	0.066	5.340	<0.01*
Platform	-0.346	0.082	-4.245	<0.01*
Residual standard error: 0.4044 on 149 degrees of freedom				
Multiple R-squared: 0.8428, Adjusted R-squared: 0.8365				
<i>F</i> -statistic: 133.2 on 6 and 149 <i>DF</i> , <i>p</i> -value: < 2.2e-16				



**Figure 5.14** Normal plot of the residual.

Table 5.18 summarizes the performance of the two models, which shows that the two models perform equally well with similar  $R$ -Squared and slightly different Adjusted  $R$ -Squared. The last model, however, is simpler with less number of predictors. As such, this model was chosen. The negative estimated coefficient of the interaction variable suggests that the variance of the scores is negatively related to changes in number of followers e.g., higher variance, lower changes in followers. For visibility richness, the positive coefficient suggests that the variance of the scores is positively related to changes in number of followers. This suggests an interesting

pattern that the public is less likely to follow celebrities who consistently promote themselves and that the visibility affordance is best utilized by being alternated. The negative coefficient of the platform variable suggests that changes in followers are higher on Instagram.

In summary, using changes in followers as a measure of audience response, I answer the fourth research question that changes in interaction richness have negative effects on the audience's growth but changes in visibility richness have positive effects on the growth.

**Table 5.18** Summary of models' performance. This shows that the last model is preferable because it is simpler but has comparable performance.

	<b>Model 4.1</b>	<b>Model 4.2</b>
Residuals are normally distributed.	Yes	Yes
Residuals are homoscedastic.	Yes	Yes
No multicollinearity.	Yes	Yes
<i>R</i> -Squared, Adjusted <i>R</i> -Squared.	0.8428, 0.8354	0.8428, 0.8356

## 5.4. Summary

The *t*-tests presented in the previous sections show that celebrities in my samples employed different microcelebrity strategies on different platforms. On Instagram, they created more identity-rich posts but fewer interaction- and visibility-rich posts. When looking at the differences by celebrity types, the analysis shows that their practices are essentially identical in all but the interaction dimension. However, the difference is only observed on Instagram. Specifically, Internet celebrities have significantly higher interaction richness than mainstream celebrities do.

Using regression, I found relationships between engagement scores and richness measures. The first model (Table 5.6) shows that the engagement scores tend to be higher with identity-rich and interaction-rich posts, but lower with visibility-rich posts. Looking at the consistency of the practices, I constructed another model (Table 5.11) to explain the relationships between mean engagement scores and variance richness scores. Among the three richness dimensions, the only significant effect yielded by the model is the variance of interaction scores. Specifically, engagement scores are negatively correlated with the variance of interaction richness scores. As score variances increase, engagement scores decrease, which suggests that audience members are less engaged with the celebrities who exhibited inconsistent interactional strategies.

The other proxy of public response is changes in numbers of followers. The model presented in Table 5.14 shows that none of the richness score had a significant effect on the changes of followers. However, the variances of interaction and visibility richness scores do impact followership (Table 5.17). Specifically, the changes in numbers of followers are negatively correlated to the variance of interaction scores, but positively correlated to the



variance in visibility scores. That is, the public might be less likely to follow celebrities who had an inconsistent interactional strategy but more likely to follow those who did not promote themselves (or the account) all the time.

The table below summarizes the questions, methods and results from this phase of the dissertation. Drawing on the results, I propose three thesis statements regarding how celebrities and audiences co-construct the environment on social media, each of the statements is strengthened and explained by the follow-up qualitative study.

**Table 5.19** Summary of research questions, methods, results and proposed statements.

Research Questions	Analytical Methods	Results	Thesis Statements
<p><b>RQ1:</b> Along the core practices, how do celebrities engage in different activities on different social media platforms?</p>	<p>Paired <i>t</i>-tests</p>	<p>Celebrities use Instagram for constructing and/or expressing their <i>identity</i> more often than on Twitter. On the other hand, they use Twitter for <i>interacting</i> with others and promoting their <i>visibility</i> more often.</p>	<p>The practices of microcelebrity differ by <i>platforms</i> and <i>celebrity types</i>.</p>
<p><b>RQ2:</b> How are the practices similar and/or different amongst mainstream and Internet celebrities engaging in microcelebrity?</p>	<p><i>t</i>-tests</p>	<p>Internet celebrities use Instagram for <i>interacting</i> with others more often than the mainstream celebrities do.</p>	

Research Questions	Analytical Methods	Results	Arguments
<p><b>RQ3:</b> How do the audiences respond to different types of strategies when controlled for celebrity types?</p>	<p>Regression analysis</p>	<p>Audiences are more likely to engage with <i>identity-rich</i> and <i>interaction-rich</i> posts, but less likely with <i>visibility-rich</i> posts. The richness, however, does not affect follow and unfollow decisions</p>	<p>Microcelebrity strategies are essential to <i>maintain</i> but <i>not to grow</i> an audience.</p>
<p><b>RQ4:</b> How do the audiences respond to the changes in microcelebrity strategies?</p>	<p>Regression analysis</p>	<p>Audiences are less likely to engage with, and follow, celebrities who exhibited <i>inconsistent interactional</i> strategies. They, however, are more likely to follow celebrities who alternated the <i>visibility promotion</i> activities.</p>	<p>Consistency in microcelebrity strategies is essential to <i>grow</i> and <i>maintain</i> an audience.</p>

## CHAPTER 6

### QUALITATIVE METHODS

The second phase of the study employs a qualitative approach to explain and validate the findings from the quantitative analyses. Whereas the quantitative component of the study illustrates the relationships between microcelebrity strategies, as measured by the framework, and an audience's responses using two proxies (i.e., followership and audience engagement); using qualitative methods allows me a better understanding of the reasons and logics behind such responses. Specifically, I collected the data by conducting interviews with the celebrities' audience members. This chapter presents the design and methodology for the qualitative analysis. The recruitment, interview protocol and coding methods are also presented in this chapter.

#### 6.1 Study Design

This component of the study aimed to address **RQ5**: *Why do the audiences respond to celebrities the ways they do?* through the examinations of the three thesis statements: 1) *The practices of microcelebrity differ by platforms and celebrity types.*; 2) *Microcelebrity strategies are essential to maintain but not to grow an audience*; and 3) *Consistency in microcelebrity strategies is essential to grow and maintain an audience*. Specifically, I used qualitative methods to gain understanding into the reasons and logics of celebrity-fans relationships from the audience perspective. If we think of microcelebrity as a performance, understanding audiences is

essential as Goffman (1959) suggests, a performance is “molded and modified to fit into the understanding and expectations of the society” (p. 35). Researchers have expanded Goffman’s argument to argue that audience members or fans play a role in co-constructing celebrity performance and the media environment within which celebrities operate (Papacharissi & de Fatima Oliveira, 2012; Senft, 2008; Thrall et al., 2008; Usher, 2015). Therefore, understanding expectations and behaviors of fans is relevant not only as a means to enhance the practice outcome and sustain promotional activity (Usher, 2015), but also as a contribution to our understandings about contemporary celebrity-fans relationships mediated by social media.

This phase of the study aims at providing an insight into the relationships between how celebrities utilize social media affordances and an audience’s responses from the perspective of audience members. The significance of this phase is twofold. First, the findings helped validate and explain the results from the prior statistical inference. Second, this study provided greater insights into audience’s expectations which were captured by the framework but could only be fully understood from the perspective of the audience.

### *6.1.1 Participant Recruitment*

My inquiry focuses on an audience’s responses to celebrities on social media. The data for this study were collected using semi-structured interviews with social media users who recently interacted with celebrities. The participants were offered \$15 for a completed interview. The recruiting letter is presented in the Appendix. In total, I conducted fifteen interviews. The interviews ranged from 54 minutes to 1 hour and 15 minutes, with an average of 58.4 minutes.

To ensure the interviews answer my questions regarding fan-celebrity experience, I defined two inclusion criteria for the study's population. That is, individuals must be qualified to participate in my study to ensure they have first-hand experience interacting with celebrities on social media (Robinson, 2014). The eligibility criteria for participants are simple. First, candidates must be actively using, and following celebrities on Twitter and Instagram. Second, they must have interacted with celebrities on Twitter or Instagram such as by commenting or liking. I employed two methods of recruitment: through the Direct Message feature of Twitter<sup>9</sup> and Instagram<sup>10</sup>, and an open platform for recruiting interview participants, [www.userinterviews.com](http://www.userinterviews.com). The combination of the two methods enhanced the generalizability of the study by capturing a wide range of perspectives regarding the experiences interacting with celebrities on social media. The approach also provided heterogeneous samples by strategically accessing multiple networks of participants (Penrod, Preston, Cain, & Starks, 2003).

To make use of the Direct Message features, a list of potential informants for recruitment was generated using lists of users who liked/retweeted or replied/commented on celebrities' posts the most in the last month of my datasets. The list of candidates was a collection of the top 500 Twitter and Instagram accounts, all of whom were contacted via the Direct Message feature of Twitter and Instagram. This feature allows users to send a private message to any account. From the total of 500 invitations, I successfully scheduled and conducted seven interview sessions. It should be noted that the low response rate might be due to the self-selection bias (Robinson, 2014). More specifically, the individuals who consented to participate might exhibit a

---

<sup>9</sup> <https://about.twitter.com/directmessages>

<sup>10</sup> <https://help.instagram.com/1750528395229662/>

special characteristic – for example, they might be more interested in the topic than others and/or more sensitive to the monetary incentives.

The other recruitment method is via [www.userinterviews.com](http://www.userinterviews.com) – an open platform for recruiting interview participants. The platform is similar to Amazon Mechanical Turk but specifically designed for qualitative research projects. Through the platform, researchers can recruit participants by posting project’s descriptions, eligibility criteria and pre-screen questions. For my study, I created a project titled *Celeb-Fans on Social Media*, provided a description (similar to the recruiting letter, noted in the Appendix) and included four pre-screen questions listed below.

- Do you have Twitter and Instagram account? Please provide your usernames (no information will be obtained from your accounts).
- Do you follow any celebrities or famous people on Twitter and Instagram?
- Did you recently interact with celebrities or famous people on Twitter and/or Instagram (e.g., by liking or commenting)? Any interaction counts whether or not you received a response.
- Please tell us about one time you have commented on or liked a celebrity or famous person's Instagram or Twitter.

The first question asks for usernames for a validation purpose. Specifically, I manually looked up the usernames to ensure they really had accounts on Twitter and Instagram. The last question asks for an experience when they interacted with a celebrity. In total, I got 117 responses with eligible qualifications i.e., have accounts on both platforms and interacted with

celebrities on Twitter/Instagram. I randomly selected and reached out to eight candidate participants to schedule an interview, all of whom agreed to participate in my study.

### *6.1.2 Interviews*

The interview protocol was designed based on results from the quantitative analyses of social media data. This study was a supplementary data set and analysis for the results from the framework, and thus strengthens the interpretations.

All fifteen one-on-one interviews were conducted on, and recorded with Adobe Connect, a web conferencing software. The lengths of interviews varied from 54 minutes to 1 hour and 15 minutes. The interviews were semi-structured; they were guided by a set of open-ended questions and follow-up questions to draw out more information from informants. The sequences of questions, attention and time spent on each topic were altered as appropriate. My informants were allowed to explore new ideas, and relate to their uses of Twitter and Instagram when interacting with celebrities on their own terms (Robson & McCartan, 2016). The interview protocol comprised opinion/belief questions (Kathwohl, 2009) and included a heading (study title, date/time and interviewer/interviewee), opening statements (release form, approximate length, purpose of research and methods of dissemination results), key questions, probes (i.e., follow-up and clarifying questions), prompts (i.e., range of possible answers), and transitional messages to move between key questions (Robson & McCartan, 2016; Strauss & Corbin, 1990). The interview protocol was defined and developed after analyzing the results from the quantitative studies. The protocol is presented in the Appendix and explained below.



The interview protocol began with a couple of demographic questions like age and education, then asked participants to talk about themselves in general e.g., life experience and work. These questions served as ice breaker but also were helpful in providing background information about participants. The protocol also included a set of questions regarding the uses of social media. Specifically, I asked for the numbers of accounts they were following, being followed and age of their accounts. I further asked for the reason they joined the platforms. These questions were aimed to situate the participants in the context of Instagram and Twitter. Then, I specifically asked for the main reasons they used Twitter and Instagram for, and asked them to describe how they usually used the platforms – for example, did they frequently post or interact with others, or lean towards passive uses.

Then, I proceeded to the context of the study by asking for a definition of celebrity. I delved deeper into the context by asking about Internet celebrities, and how they were different from mainstream celebrities. After this set of questions, I emphasized that for the rest of the interview the term *celebrity* would include both mainstream and Internet celebrities. It is important to note that this clarification needed to be frequently restated throughout the interview.

With the established context of celebrity, the protocol began to investigate the use of social media by asking for the number of celebrity accounts (percentage-wise) the participants were following, asking them to split those into mainstream and Internet celebrity accounts. To get them really thinking about celebrity accounts, I asked for examples of mainstream and Internet celebrity accounts before further interviewing. Then, I asked for reasons and expectations from following celebrities on social media and whether or not the expectations were different for different types of celebrities. I also asked why (or why not) they followed the same celebrity accounts on both platforms to gain deeper understanding into their expectations and

perceptions of how celebrities utilized different platforms. Then, I began investigating their decisions to follow and unfollow celebrity accounts and what propelled them. I also asked about their decisions to engage with celebrities by urging them to talk about different forms of engagement by using a simple term like *interaction with celebrities*. Practically, I asked them to describe what they typically do to interact with a celebrity on social media and why they choose to do so. I ended this section by asking the participants to talk about limitations of the platforms and what they might want to change about the platforms in the context of celebrity-fans relationship.

The protocol then shifted to investigate participants' perception on how celebrities utilized the identity affordance of social media. I framed the questions using words like *personas, character* or descriptively like *celebrities presenting themselves*. Specifically, I asked the participants to describe how they saw celebrities expressing themselves using the features (artifacts) of social media and if there were any differences between celebrity types. Then, I inquired how each of the identified features helped them imagine or create a picture of celebrity's character or personality, and if/how such features played a role in their decision to engage with a celebrity. I also asked the participants whether or not they would unfollow the celebrities if they stop engaging in actions related to presenting their persona on social media.

The next set of questions concerned participants' perception on how celebrities utilized the interaction affordance of social media. Specifically, I asked them to talk about social media features (artifacts) celebrities used to interact with fans, how such actions affected their decision to engage with the posts, and if the effects differed by celebrity types. I also asked if they ever did anything in an attempt to get a response from a celebrity, if so, how it went and how they felt

about that. I further questioned how they would feel if the celebrities stopped interacting with the public on social media.

Another set of questions concerned participants' perception on how celebrities utilized the visibility affordance of social media. I began by asking how they typically find or get to know the celebrity accounts they are following on social media. I further asked what social media features the celebrities could use to promote their accounts to gain more followers or expand the audience. Then, I asked if the use of such features affected their decision to engage with a celebrity and if the effects differed by celebrity types.

The last set of questions was about the perception of fan-celebrity community on social media. I began by asking the participants to describe the fan communities on social media, what the communities looked like and if they felt they were part of the communities. Then, I asked the participants to describe the actions they typically took to interact with other members in the communities, or at least the actions they saw others used to interact. I ended this section by asking the participants to compare the dynamics of fan-celebrity relationship in pre- and post-social media era.

All interviews were audio-recorded and subsequently transcribed by myself for further analysis. The coding was done on TAMS Analyzer or Text Analysis Markup System, a software for coding and extraction for qualitative studies.

## **6.2 Qualitative Coding**

The coding process employed an approach that gradually allowed themes to emerge as realized through information reduction, conceptualization, elaboration and relating (Strauss &

Corbin, 1990). That is, the information was reduced to certain patterns or categories to develop a coding scheme using a template analysis approach (Robson & McCartan, 2016). As such, I established themes through a process of collapsing, challenging, and merging codes through axial coding, which led to the themes that helped me interpret the relationships discovered in the previous phase.

To ensure the accuracy and validity of the analysis, I performed a *member-check* during the interview process by restating, summarizing the information and questioning the informants. I also performed a post-analysis member-check by asking the informants to affirm that my interpretations reflected their experience and views (Robson & McCartan, 2016; Yanow & Schwartz-Shea, 2015).

The first cycle coding organized the raw transcriptions into initial codes. Then, the transcripts were segmented into the relevant codes if applicable; otherwise a new code was emerged. Once the initial coding scheme was finalized, the transcripts were re-coded. Specifically, I used a combination of in-vivo and structural codes to preserve the languages of this specific group of participants while maintaining the themes which ran through the interview protocol. I identified codes related to the context of the study e.g., reasons for using social media and definitions of celebrities. Examples of the initial codes include ‘connect with family’, ‘connect with people I know in real life’, ‘person who is famous’, ‘recognizable by many people’, ‘a quirky person’, ‘talents or some unusual quirks’, ‘available but not approachable’ and ‘both types of celebrity are pretty much the same to me’.

I also identified and organized codes by the core practices i.e., responses related to identity construction (e.g., ‘show what they really are through pictures’, ‘talk about themselves’, ‘not expect Internet celebrities to do this more often’, ‘celebrities show openness’, ‘being part of

their circle’), interaction (e.g., ‘respond to us’, ‘ask questions’, ‘easier to do on Instagram’, ‘Internet celeb needs to work harder’, ‘the key is to maintain the interactions’) and promoting visibility (e.g., ‘hashtags are really helpful’, ‘post a lot and get themselves out there’, ‘annoying when they do that a lot’, ‘how effortless they are’).

I re-analyzed the initial codes to explore the interrelationships across multiple codes to develop a coherent synthesis of the data using an axial coding approach. That is, I related the codes to sub-categories, defined their labels and locations to address the “if, when, how, and why” questions (Charmaz, 2006, p. 60).

**Table 6.1** Axial codes and descriptions.

<b>Axial codes</b>	<b>Definition</b>
<p>1. My social media uses</p> <ul style="list-style-type: none"> <li>- For connecting with friends and family or people I know in real life.</li> <li>- For entertaining or passing time.</li> <li>- For obtaining information, news and updating trends.</li> <li>- For sharing or informing others of personal updates.</li> </ul>	<p>This category explains general information about participants’ social media behaviors, comprising of four codes as the main reasons participants are using social media for.</p> <p>Examples: ‘see what friends are doing’, ‘photos they post’ and ‘connect with your family’.</p>
<p>2. Definitions of celebrity.</p> <ul style="list-style-type: none"> <li>- A person who is famous or recognizable by many people.</li> <li>- Someone with audiences or groups of people interested in his/her private life.</li> </ul>	<p>This category explains participants’ perceptions of the celebrity concept, comprising of three codes.</p> <p>Examples: ‘famous for whatever reason’. ‘a public life.’, ‘recognizable by many people’ and ‘a quirky person’.</p>

<ul style="list-style-type: none"> <li>- A charismatic person or someone with talents or unusual abilities.</li> </ul>	
<p>3. Celebrities' social media uses.</p> <ul style="list-style-type: none"> <li>- To bypass mainstream media, organizations or institutions.</li> <li>- To increase accessibility for their fans.</li> <li>- To share details or updates about their personal life and work.</li> <li>- To humanize or bring themselves closer to fans.</li> </ul>	<p>This category explains participants' perceptions of how celebrities use social media, comprising of four codes.</p> <p>Examples: 'humanize them' and 'make that relationship feel a little bit less formal'.</p>
<p>4. Presenting persona or <i>identity</i> on social media.</p>	
<p>4.1 Form of practices.</p> <ul style="list-style-type: none"> <li>- Celebrities present themselves through textual data.</li> <li>- Celebrities present themselves through their photos.</li> </ul>	<p>This sub-category identifies social media artifacts for celebrities to appropriate the identity affordance.</p> <p>Examples: 'express yourself through photo'.</p>
<p>4.2 Twitter vs. Instagram.</p>	<p>This sub-category explains if one platform is more appropriate than the other for celebrities to appropriate the identity affordance.</p> <p>Examples: 'tougher to do on Twitter'.</p>
<p>4.3 Mainstream vs. Internet celebrities.</p>	<p>This sub-category explains if different types of celebrity appropriate the identity affordance similarly or differently.</p> <p>Examples: 'Internet celebrities are more involved presenting themselves'.</p>
<p>4.4 Following decisions.</p>	<p>This sub-category explains the effects of the practices on follow/unfollow decisions.</p> <p>Examples: 'they don't owe me any sort of life updates'.</p>

4.5 Engagement decisions.	This sub-category explains the effects of the practices on engagement decisions.
5. Fan <i>interaction</i> on social media.	
<p>5.1 Form of practices</p> <ul style="list-style-type: none"> <li>- Celebrities interact with fans by responding to, or recognizing fans' service.</li> <li>- Celebrities interact with fans by asking questions or asking for comments/feedback.</li> </ul>	<p>This sub-category identifies social media artifacts for celebrities to appropriate the interaction affordance.</p> <p>Examples: 'ask questions' and 'ask for comments'.</p>
5.2 Twitter vs. Instagram.	<p>This sub-category explains if one platform is more appropriate than the other for celebrities to appropriate the interaction affordance.</p> <p>Examples: 'happens on Twitter more often'.</p>
5.3 Mainstream vs. Internet celebrities.	<p>This sub-category explains if different types of celebrity appropriate the interaction affordance similarly or differently.</p> <p>Examples: 'Internet celeb comment or responds to the fans more than regular celeb'.</p>
5.4 Following decisions.	<p>This sub-category explains the effects of the practices on follow/unfollow decisions.</p> <p>Examples: 'the key is to maintain the interactions'.</p>
5.5 Engagement decisions.	<p>This sub-category explains the effects of the practices on engagement decisions.</p> <p>Examples: 'check if he responses to a lot of posts'.</p>
6. Promoting <i>visibility</i> beyond the existing fan base on social media;	
<p>6.1 Form of practices.</p> <ul style="list-style-type: none"> <li>- Celebrities promote their visibility by using hashtags.</li> </ul>	<p>This sub-category identifies social media artifacts for celebrities to appropriate the visibility affordance.</p> <p>Examples: 'hashtags increase post's visibility' and 'get you the audience that you would not normally have'.</p>

- Celebrities promote their visibility by being active.	
6.2 Twitter vs. Instagram.	This sub-category explains if one platform is more appropriate than the other for celebrities to appropriate the visibility affordance.
6.3 Mainstream vs. Internet celebrities.	This sub-category explains if different types of celebrity appropriate the visibility affordance similarly or differently.
6.4 Following decisions.	This sub-category explains the effects of the practices on follow/unfollow decisions.
6.5 Engagement decisions.	This sub-category explains the effects of the practices on engagement decisions.  Examples: ‘too many unnecessary hashtags’ and ‘annoying’.

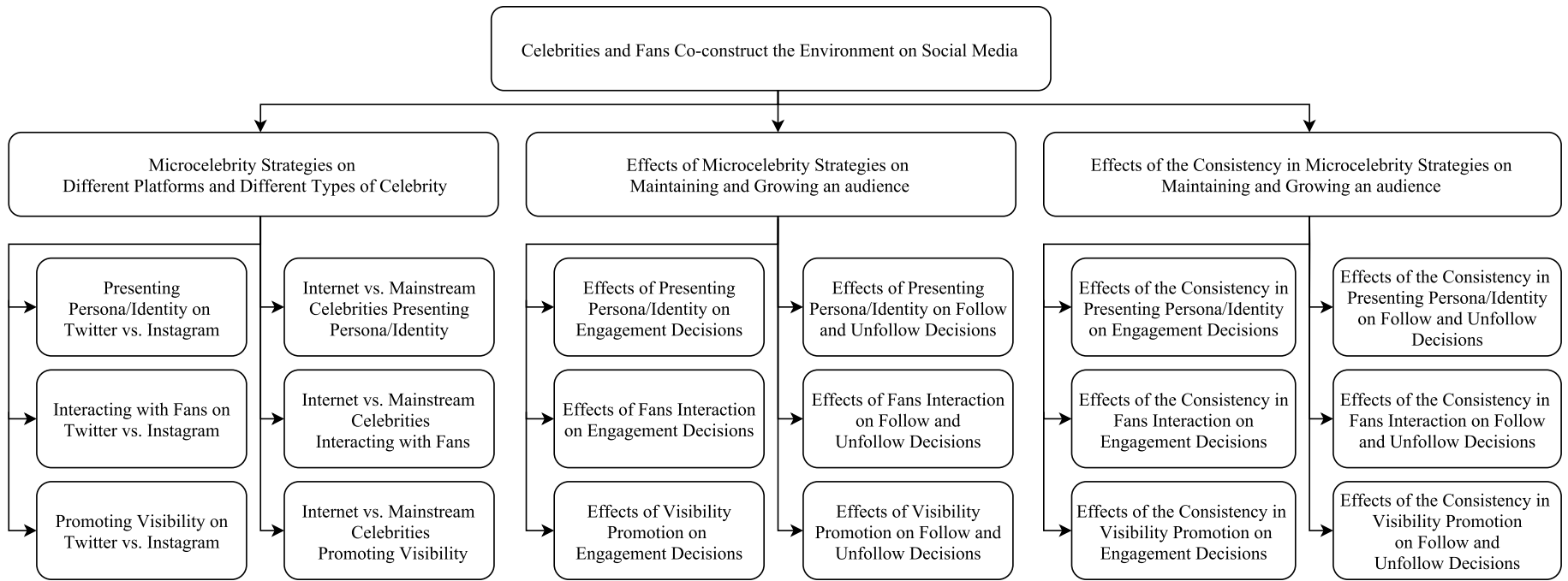
With the axial codes, I used a focused coding approach to identify the most frequent and significant codes that could generate the most analytic traction by constantly comparing codes against codes. Lastly, I used a theoretical coding to identify the primary umbrella *theme* or a core thread that ran through the data to systematically link categories together. The theme for this work is *the process of co-constructing microcelebrity performance by celebrities and fans, mediated by social media*. This theme captures how celebrities and their fans utilize social media affordances to co-construct the performance and media environment by looking at celebrities’ activities (focused code # 1) and the responses from the audiences (focused codes # 2-3).



**Table 6.2** Coding process showing a transition from initial codes, axial codes. focused codes and theoretical codes.

Examples of initial Codes	Axial Codes	Focused Codes	Theoretical Codes
Express yourself through photo, tougher to do on Twitter, happens on Twitter more often	Presenting persona/identity, interacting with fans and promoting visibility, on Twitter vs. Instagram	Microcelebrity strategies on different platforms and different types of celebrities	The co-construction of microcelebrity performance by celebrities and fans, mediated by social media
Internet celebrities are more involved presenting themselves, Internet celeb comment or responds to the fans more than regular celeb	Internet vs. Mainstream celebrities: presenting persona/identity, interacting with fans and promoting visibility		
Check if they response to a lot of posts, too many unnecessary hashtags, annoying	Effects on engagement decisions	Effects of microcelebrity strategies on maintaining and growing an audience	
They don't owe me any sort of life updates	Effects on follow/unfollow decisions		
The more I see celeb interacting with people the more I want to comment, something they truly do care with and that they like to respond	Effects on engagement decisions	Effects of the consistency in microcelebrity strategies on maintaining and growing an audience	
The key is to maintain the interactions	Effects on follow/unfollow decisions		

The final codebook is explained below and presented in Figure 6.1. More details and examples are presented in the Appendix. Under the primary theme, the co-construction of microcelebrity performance mediated by social media, is the three focused codes. The first code concerns the similarities and differences of microcelebrity practices on different platforms and different celebrity types. This code consists of six sub-codes: the first three sub-codes are related to each of the core microcelebrity practices on Twitter vs. Instagram; the other three sub-codes are related to the core practices engaged by Internet vs. Mainstream celebrities. The second focused code looks at the effects of the practices on an audience's decisions to engage with celebrities as well as decisions to follow and unfollow celebrities' accounts. This code consists of six sub-codes: the first three sub-codes are the effects of the core practices on engagement decisions, the others are the effects on follow and unfollow decisions. The last focused code concerns the consistency in microcelebrity practices and comprises six sub-codes. The first three sub-codes are the effects of the consistency of the core practices on engagement decisions, the others are the effects on follow and unfollow decisions.



**Figure 6.1** The Structure of the Codebook for Analyzing Interviews Data.

### **6.3 Summary**

In this chapter, I presented the design and methodology of the qualitative phase of my research. I presented how the participants were recruited and the methodology for the data coding as well as the final codebook. In the next chapter I present the results and findings from the qualitative study and discuss how they support the results from the quantitative phase.

## **CHAPTER 7**

### **QUALITATIVE RESULTS AND FINDINGS**

As noted earlier, I conducted a qualitative study by collecting the data from the interviews with audience members. This analysis was designed to be a supplementary study to explain and validate the findings from the quantitative analyses within the perspective of the co-construction of microcelebrity performance by celebrities and their fans. This chapter presents a broad overview of my informants and organizes the coding results by their themes. The chapter closes with a discussion on how the results support the findings from the quantitative study through the examinations of the thesis statements proposed earlier.

#### **7.1 Findings**

To give a broader view about my informants, ten of them are female and the majority have a Bachelor's degree (one has a graduate degree and one is still in college). The average age is 28.33 with a minimum of 18 and maximum of 45 years old. In terms of Instagram profiles, most of them have had the account since 2013 with an average number of followers of 587.72 and an average of 512.07 followings (i.e., Twitter's Friends). For Twitter profiles, most of them have had the account since 2008 with an average number of followers of 1,033.6 and an average of 8,18.89 followings. Most of them described themselves as a Poster (i.e., they frequently post). The average number of celebrity accounts they followed was 38% with a minimum of 10% and maximum of 80%.

I began each interview with general questions about their social media uses. The top two reasons for using social media are *connecting with friends, family or people I know in real life* and *obtaining information, news and updating trends*. I also asked for a definition of celebrity, which most of them explained as *a person who is famous or recognizable by many people*. I also delved deeper into this by asking what they would think of the concept of Internet celebrity and the differences between mainstream and Internet celebrities. The informants consistently did not report many differences except for the source of the fame. Specifically, they also thought of Internet celebrities as a famous and unapproachable figure but the difference was that Internet celebrities were famous from what they did on social media. For example, P3 explained “*They are all celebrities to me*” together with P11 who reported “*More and more I see the line between the two [Internet and mainstream celebrities] graying*” and P4 who further commented “*I think they are the same but Internet celebrities have become famous on the Internet for a reason*”.

I also asked for some examples of celebrity accounts (of both types) to get the informants to think about such accounts before further questioning. For mainstream celebrities, the most common examples are actors/actresses and musicians. The examples of Internet celebrities are more diverse including comedians, beauty bloggers and fashionistas.

In the following sub-sections, I present the findings organized by the themes of audience’s opinions and belief within the data identified earlier.

### *7.1.1 Microcelebrity Practices on Different Platforms and Different Celebrity Types*

Under the primary theme of microcelebrity performance co-constructed by celebrities and their fans, the analyses first examined the practices conducted by celebrities. When asked about perceptions of how celebrities presented their identity online, I framed the question using words like *personas* and even more simple phrases like *presenting themselves or their character/persona*. Most of the informants instantly reported Instagram as a platform for presenting one's identity. Yet they could also explain how celebrities presented their personas when urged to talk about the practices on Twitter. The ways they saw celebrities presenting their identity were: posting photos, providing captions and talking about themselves. A few informants tended to think that Internet celebrities need to work harder at presenting themselves than mainstream celebrities because the Internet was their primary channel.

Next, I inquired about the interaction work of celebrities on social media. Very few informants tended to think of the interactivity on Instagram; most of them instantly talked about Twitter and explained that interactions on Instagram were quite invisible in a sense that you could not really see if the interactions took place. Interactions could be responding to comments and questions or recognizing fans' service. Celebrities also interacted with fans by asking questions addressing the audiences or asking for comments or feedback. P1 explained:

*“They'll ask questions or comment to say thank to people that say things or just depends on how engaging they feel like being.”*

When asked about the interactivity of mainstream and Internet celebrities, a majority of informants thought Internet celebrities were more interactive with fans. For example, P7 and P8 explained, respectively:

*“I’d say Internet celeb are more likely to comment back more frequently.” and “They [mainstream celebrities] don’t often interact with their audience using the Internet ... but Internet celebrities seem to have more hands-on and connect to the people.”*

Another set of questions inquired about the ways celebrities promote visibility beyond the existing fan base. When asked how celebrities promoted their accounts on Twitter and Instagram, the informants seemed to think about purchasing ads although they never saw any celebrities using them. One informant suggested posting at a certain time or day could gain higher visibility than at others. My informants also mentioned hashtags as a way to increase visibility although only a few informants had found new celebrity accounts to follow through the uses of hashtags. P15 explained:

*“A lot of Internet celeb, add like a billion of hashtags to the end [of the post] so they show up in more people’s searches.”*

Hashtags are also helpful for celebrities to gain access to the new group of audiences as P8 explained:

*“I do think hashtags are pretty helpful because they definitely increase post’s visibility and get you the audience that you would not normally have.”*

Together, this analysis suggests that the ways celebrities engaged in microcelebrity practices differ by social media platforms and that the practices differ among celebrity types.



### *7.1.2 The Effects of Microcelebrity Strategies on Maintaining and Growing an Audience*

The other perspective to look at the co-construction of microcelebrity performance is through the responses from audiences. When we talked about the ways in which celebrities presented themselves to an audience, I asked how such practices affected audience's decision to follow or unfollow celebrity accounts. Many informants were quite selective about their followings as P7 explained,

*“I think I follow a lot of people but I think I'm pretty choosy about who I follow.”*

However, most informants reported they did not really go over the existing posts before making a following decision. They typically knew or heard of the persons from friends, other social networks or media, and would follow the accounts only if they wanted to learn more about them. Neither was the unfollow decision driven by the identity practices. For example, P1 reported:

*“I've seen someone else tweeting them or following them and it shows up on my feed and I chose to follow based on that.”*

I also asked about the decisions to engage with celebrities. The identity factor seems to play a more important role in the decision to engage, where a majority reported they were more likely to respond to posts when celebrities talked about themselves either by posting pictures or giving access to daily life. P12 reported:

*“When I see a picture of Justin Bieber, I’d go crazy. And I like to click like because of a particular person that is on there. When I see Justin, I click like right away. You don’t have to do anything, just post the face.”*

Similar to the identity practices, the interactions did not play an important role in the following decision, where informants explained it was not even their expectation to see celebrities interacting with fans. Although interactions were not expected, it would be very exciting when celebrities did interact with fans. Such interactions are really appreciated as P8 said:

*“It means a lot to me when I see someone who I look up to or whose work I like, when I see them interacting with like casual people like me like fans, like people who are not as creative but they are still trying.”*

In fact, the interactions could encourage the fans to be more engaging. Most of my informants explained they would be more likely to engage with celebrities who have interacted with fans before, P5 reported:

*“I guess if I felt the need to interact with the post like if I have a comment, I might then go back and do something and check like if he [a celebrity] responds to a lot of posts, I feel ok about commenting on it too.”*

Although the informants reported using hashtags as a way to boost up visibility and gain access to the new groups of audiences, using too many hashtags could come off as *a little taggy* and made them less likely to engage with the posts. P14 and P8 reported, respectively:

*“I think sometimes it could come off as a little taggy if you have too many hashtags in your post.” and “If I think there are too many unnecessary hashtags, there's a slight chance that I'll just skip it just because it is kind of annoying.”*

This analysis suggests that audiences respond differently to different microcelebrity strategies and that we might expect celebrities to adjust their practices to suit with an audience's expectations. That is, the performance of microcelebrity is shaped by the audience.

### *7.1.3 The Effects of the Consistency in Microcelebrity Strategies on Maintaining and Growing an Audience*

Within the perspective of responses by audiences as part of the co-construction of microcelebrity performance, this analysis looked at how the audiences responded to how consistent celebrities engaged in the practices. Some informants suggested many celebrities had fewer self-promotion posts and more sponsored posts once they reached a certain level of audience. However, this was understandable and did not make them unfollow the accounts.

Although interactivity had no effects on the follow and unfollow decision, interestingly some informants reported the consistency of interactions was a key to the unfollow decision. Specifically, they expected celebrities to maintain the interaction level with fans as P4 explained:

*“If they never interact with people before, I'll be like fine that person doesn't really interact. But if they used to and now they don't, I would unfollow them.”*

During our conversation around celebrities promoting their visibility, a majority of informants explained such practices were sometimes annoying and showed an obvious attempt of

attention-seeking. This means that although the visibility promotion practices are helpful for expanding the audience, they might dissatisfy the existing audience thus celebrities should find a balance. P11 reported:

*“People [celebrities] are actually posting about their life and sprinkle in perhaps giveaways or challenges to attract more followers, I think they need to be balanced”*

Interestingly, one of my informants, P15, noticed an attempt to conceal the attention-seeking behaviors on Instagram by including hashtags in a comment rather than directly in a post. In this way, the hashtags would not show up but be hidden under other more recent comments.

*“When a celeb does that [using hashtags], that is definitely annoying, ..., but they usually use it as a comment so it does not really annoy me because you cannot see unless you click on it [the comment section].”*

This analysis suggests that the consistency in how celebrities engaged in the practices plays an important role in an audience’s impression towards celebrities. The results also suggest an evidence that celebrities are aware of an audience’s expectations and so adjust their strategies to fit with the expectations. Therefore, the results are in well support of the argument claiming that microcelebrity performance is co-constructed by celebrities and their fans.

## **7.2 Summary**

This chapter presented the findings from interviews with the audience members of celebrities. With the protocol design based on quantitative results, the interviews helped interpret

and strengthen the findings from the previous phase through the examinations of the three thesis statements. Specifically, the first statement: *The practices of microcelebrity differ by platforms and celebrity types*, was well supported by the interviews explaining Instagram was more suitable for celebrities presenting themselves and the informants who considered pictures as the best way to gain insights into celebrities' life. Moreover, the interviews also provided an explanation that Internet celebrities needed to work harder to maintain their fans especially on the Internet, probably because that was where their fan base was.

The interviews also provided evidence to corroborate the second statement: *Microcelebrity strategies are essential to maintain but not to grow an audience*. My informants explained that they were more likely to respond to posts when celebrities talked about themselves either by posting pictures or giving access to daily life. Although my informants did realize that celebrities might not read through all the comments or responses they got on social media, they would still respond when celebrities asked questions or simply asked for feedback or opinions. When asked about the accounts they chose to follow, my informants explained that although they were quite selective about the accounts they followed, they rarely looked at the posts of celebrities but based their decision on what they have heard about the celebrities. For example, the accounts might be suggested by their friends. This explains the last part of the statement that microcelebrity strategies have no effect on growing an audience.

The last argument: *Consistency in microcelebrity strategies is essential to grow and maintain an audience*, was also confirmed. Interestingly, fans understood if a celebrity did not want to interact much; my informants explained it was not even their expectation to see celebrities interacting with fans. My informants, however, seemed to expect the celebrities be consistent about the ways they interacted with fans. Specifically, they expected celebrities to

maintain the interaction level with fans. My informants also commented on the promotional activities. Although they noted that it was understandable, especially for the Internet celebrities, to be engaging in a lot of self-promotion activities, consistently doing so could be annoying. This suggests that audience members preferred their celebrities to strategically engage in self-promotion activities by frequently alternating their visibility-promotion practices.

Together, this reflects the nature of the *mediated* microcelebrity performance as a co-construction process of celebrities and their fans. Social media have complicated the dynamic of celebrity-fan relationships by creating a new set of expectations – for example, audiences expect more intimate relationship from a celebrity, or expect them to be more interactive on social media. On the one hand, celebrities utilize multiple social media platforms to manage parasocial relationships by mixing the richness as appropriate. On the other hand, fans are given the feedback channels which help shape the performance of celebrities by signaling what practices are cherished and what are not. This shows that the mediated performance in the age of social media gradually gives more power to fans in the celebrity management model, which was once a highly controlled and regulated institutional model (Marwick & boyd, 2011; Turner, 2013).

## CHAPTER 8

### DISCUSSION

In this dissertation, I examine the practices of microcelebrity on social media through the perspective of affordances or action possibilities (Gibson, 2014a). Microcelebrity is a set of self-presentation techniques through the uses of technologies like social media (Senft, 2008). The theory explains that people construct their public persona as a *commodity sign* or product to be consumed by others (Hearn, 2008), using strategic intimacy to appeal to followers (Senft, 2008), and regarding their audience as fans (Marwick & boyd, 2011). With social media, celebrities can bypass the mainstream media and interact and communicate with the public directly. As a result, they have more control over the presentation of their persona and the relationship they have with fans (Turner, 2013).

This chapter presents discussions around the methods and findings of this work in five aspects. First, I present a discussion of the research method. Previous studies on microcelebrity use either a qualitative or quantitative method. This study employs a novel mixed-method research design consisting of both quantitative and qualitative methods. I justify the method as an appropriate choice for the study, and show how it provides a profound insight into the practices of microcelebrity.

The second discussion is on the development of the framework and the notion of *affordance richness*. I explain how the framework and the notion can be adopted in future studies. Third, I present a discussion around the findings and what they mean when considered

through the theories and literature discussed in Chapter 2. The findings related to each of the core microcelebrity practices are separately discussed.

Then, I present a discussion of microcelebrity as a learned technique. I discuss how the knowledge and implications from the study would be useful for people seeking attention or celebrity status online. The last aspect of the discussion is on limitations and possibilities for future studies.

## **8.1 Research Methods**

As noted in Chapter 3, microcelebrity studies have adopted a variety of research approaches including qualitative and quantitative analyses. A qualitative approach could provide meaningful insights from the practitioner's point of view (Abidin, 2015; Bennett, 2014; Huba, 2013; Marwick, 2015a; Ward, 2016); approaching celebrities, however, can be challenging (Mavroudis & Milne, 2016). As a result, such studies are mostly small-scale studies that limit the generalizability of the study. Another group of microcelebrity studies are large scale analyses with more generalizable results through the uses of a laboratory experiment (Jung et al., 2017) and qualitative content analysis with statistical inference (Frederick et al., 2014; Golbeck et al., 2010; Hemphill et al., 2013). While each approach has its own strengths and weaknesses, the mixed-methods approach used in this work provides a comprehensive look and offers a complete picture of the results through complementary strengths and non-overlapping weaknesses (Creswell, 2013).

More specifically, this dissertation employed an explanatory mixed-methods design (Creswell, 2013), which began with quantitative analyses of social media data, followed by



qualitative study. The first phase of the study reflected my deliberate effort to obtain generalizable results through statistical inference techniques such as tests of mean equality and regression analysis. The second phase relied on a qualitative approach to provide causal explanations and to confirm and clarify the findings from the audience perspective. This is particularly important given that microcelebrity is a performance co-constructed by the practitioners, audience members and platforms like social media (Goffman, 1959, 1959; Papacharissi & de Fatima Oliveira, 2012; Senft, 2008; Thrall et al., 2008; Usher, 2015). As such, it is important to understand audiences' expectations and the reasons behind their responses to different microcelebrity strategies. The data were collected using semi-structured interviews with audience members, which were helpful in both strengthening the analyses as well as providing deeper insights into expectations and behaviors of fans. This is particularly important not only as a means to enhance the practice outcome and sustain promotional activities (Usher, 2015), but also because it contributes to our understandings about contemporary celebrity-fan relationships mediated by social media.

Through the uses of this mixed-method research approach, my study overcomes the challenges of quantitative studies (i.e., lack of deep interpretations) and qualitative studies (i.e., generalizability). Specifically, the findings from my quantitative analyses could be generalized while their interpretations were supplemented by the interviews, providing multifaceted insights from audience members.

## 8.2 The Framework and Notion of Affordance Richness

As a part of the quantitative analyses, I developed a richness framework through the conceptual lens of the theory of Affordances (Gibson, 2014a). The theory explains that affordances are the abstract high-level action possibilities that the sites afford users to perform through the uses of technology artifacts (Fayard & Weeks, 2014; Gibson, 2014a). For example, Twitter offers an editability affordance through the delete button that allows users to delete their own contributions (Treem & Leonardi, 2013). Affordances Theory has been widely adopted by scholars to study the utilities of an environment (e.g., social media sites) and explain how the environment affords users the ability to perform activities through the available artifacts.

Previous studies generally use the theory of Affordances to study an environment as a whole – for example, Wang et al. (2012) adopted the theory to study Facebook in the context of a teaching-learning environment and suggested a group of affordances based on common usage patterns. I suggest that the theory can be adopted to examine *technologies in practice*, meaning the unit of analysis becomes the daily practices of technology uses. This is because the wide range of technology artifacts offered by new technology introduces variety in usages i.e., the technology can be used in many different ways (Brinker et al., 2015).

With variety in how we use the technology, the theory does not offer a systematic way to examine how the affordances, when undertaken, enable users to engage in social actions. For Gibson, affordances always exist regardless of user perceptions or appropriations (Gibson, 2014a) – for example, a cellphone always affords archivable communication (via texting) and rapid communication (via calling) although an elderly person might not appropriate the archivable-communication affordance but a teenager might. The original theory does not provide

ways that we could use to examine the ability of this particular use of the cellphone to help users achieve their goal along two affordance dimensions (archivable and rapid communication).

To identify the common usage patterns of celebrities, I relied on the microcelebrity literature which suggests three core practices of microcelebrity: identity construction, interaction with fans and promoting visibility beyond the existing fans base. On social media, one's identity can be constructed by positioning the self in relation to others (Page, 2012), by sharing information which reflects an identity, or what impression they want others to have about them (Marwick, 2015b). Another core practice of microcelebrity is fan interaction. By interacting with fans, the practitioners can develop and sustain their audience. The last practice is to promote visibility beyond the existing fan base by engaging in the acts that promote public exposure. This is particularly important as a means to grow an audience (Turner, 2010).

I conceptualized the common patterns (core microcelebrity practices) as the set of affordances that social media offer celebrities, and developed a richness scoring framework. More specifically, the framework was developed based on the relations of the higher-level patterns of user behaviors (the core practices of microcelebrities) and the technological artifacts of social media (e.g., @mention, hashtags). I also developed the notion of *affordance richness* as the ability of a medium (i.e., tweet or Instagram post) to deliver the information necessary in affording a particular action by using artifacts of social media. Just like *media richness* is the ability of a medium to deliver rich information (Daft & Lengel, 1986), *affordance richness* measures the richness of affordances made possible by a medium (a post). For example, when a celebrity creates a tweet (a medium) with an @mention (an artifact) to interact (an affordance) with someone in the audience, I would say that the tweet is rich in interaction affordance, or that the tweet has the ability to deliver the information necessary in affording interactions through the

use of @mention artifact. In this example, I would be measuring the richness of interaction affordance. Simultaneously, I could also be measuring the richness of other affordances (i.e., identity and visibility).

To measure the affordance richness, the framework organizes the technological artifacts of social media into groupings by their relevance to affordance dimensions based on HCI literature (boyd et al., 2010; Honey & Herring, 2009; Hu et al., 2014; Kwak et al., 2011), and measures the richness of a post based on the way it is constructed. That is, the richness reflects the uses of technology artifacts within an associated affordance-artifact grouping.

I expect that this framework will provide a way for researchers to compare actors in different contexts. Finding the labels for each post allows researchers to aggregate scores in many different ways. For example, I could have found the mean/median richness scores for each of the celebrity types. This would give me, for example, one identity score for mainstream and one for Internet celebrities group. In this example, the unit of analysis would be at the celebrity-type level.

In this study, the units of analysis are post level – e.g., the regression model of the relationships between engagement scores and richness scores, and user level – for example, the paired *t*-tests used to examine if celebrities use platforms similarly or differently. With these units of analysis, I could compare not only the central behaviors among groups, but also the variance within users. This flexibility is certainly a strength of the framework, but suggests that researchers should think carefully about how different aggregations may lend themselves to a different unit of analysis.

Researchers can also leverage the notion of *affordance richness* to explain the ability of an object to afford a particular action. This could be particularly useful when the object can be used in different ways, or when the phenomenon of interest comprises multiple dimensions of practices. For example, a study of a social movement on Twitter might argue that the core practices of social movement are mobilizations and information disseminations. The study can use the concept of affordance richness to examine a corpus of tweets and divide them into two groups: one being rich in the mobilization affordance, and the other is rich in the information-dissemination affordance. With the two groups of tweets, the study can conduct further analyses – for example, to examine how different groups of users (e.g., activists and the publics) use the affordances similarly or differently.

### **8.3 Findings**

Recall the overarching question of my study is *How do celebrities use social media to grow and maintain celebrity status?* I used two proxies to measure celebrity status: audience engagement and the size of followers. With the framework, I generated the richness scores of a large set of Twitter and Instagram data from celebrities of both types (i.e., mainstream and Internet famous). Then, I answered the question with a series of statistical analyses including *t*-tests and regression modeling to explore the relationships between the richness scores and the proxies of celebrity status. Drawing on the findings, I proposed three thesis statements, each of which was confirmed and explained by the qualitative study through interviews with audience members.

My analyses show that microcelebrity performance is co-constructed by celebrities and their fans. With social media, celebrities have more control over the management of their self-presentation and relationships with fans by utilizing multiple social media platforms and mixing the affordance richness as appropriate. The findings suggest that one platform is more suitable for some practices as reflected by celebrities using them more often. My findings also indicate that differences exist between the practices of mainstream and Internet celebrities. On the other hand, fans are given the feedback channels which help shape the performance of celebrities by signaling what practices are cherished and what are not. This shows that the mediated performance in the age of social media gradually gives more power to fans in the celebrity management model, which was once a highly controlled and regulated institutional model (Marwick & boyd, 2011; Turner, 2013). Looking at the audience responses, I found that audiences were more likely to engage with the posts categorized as rich in some affordance dimensions; their follow and unfollow decisions, however, were independent of the richness scores. On the other hand, the consistency in the richness scores did have significant effects on engagement, follow and unfollow decisions. In the following section, I present a discussion of the findings, what they mean and how they relate to the literature.

### *8.3.1 Identity Construction*

My richness score analysis, along with the interviews with audience members, indicate that Instagram is more appropriate for presenting a character or showing what the person truly is. Instagram affords an identity construction, partly, through the ability to post pictures. Of course, users can post a picture of anything, but the presence of faces, one of the artifacts associated with

the identity affordance, has been identified as a driver of audience engagement on Instagram (Bakhshi et al., 2014). For the posts categorized as rich in identity affordance, the median number of faces is one with the mean of 0.8. This finding is supplemented with a response from my informants, suggesting they would be more likely to respond to the posts when celebrities talked about themselves either by posting pictures or giving access to daily life. Together, I suggest that pictures of celebrities partly afford the action possibility of constructing an identity that will be consumed by the public.

Instagram is more appropriate for identity construction not only because of the nature of the platform as a visual medium (Bakhshi et al., 2014; Jerslev & Mortensen, 2016; Marwick, 2015a), but also because of the way celebrities are utilizing the platform. Specifically, my informants reported celebrities were using the platforms differently. While Twitter is mainly used for random thoughts or small updates throughout the day, Instagram posts are more thoughtful and used for major updates. An example from my informant is, a celebrity might post pictures from an event at the end of the day on Instagram, but continuously tweet about the event throughout the day. Together, the two platforms afford celebrities the possibility to make a bigger story and create a more complete picture of the celebrities' daily lives.

In fact, this is the main reason people are following the same celebrities on both platforms, to be part of the bigger *livestreams* (Marwick, 2013). Marwick (2013) explains that *livestreaming* is an act of ongoing sharing of personal information in an attempt to create a digital portrait for one's *networked audience*. By networked audience, she means the real and potential audience for digital content, who are connected to the content creator (a celebrity) as well as to each other. With social media, the audience members can consume as well as contribute to the livestreams, all of which creates a sense of co-presence. That is, audiences

could feel as if they were with the celebrities. Some celebrities even encourage fans' contributions by asking them to share their stories or comments after an event and so on. However, the fans' contributions sometimes come with conflicts and makes impression management even more difficult, such that some celebrities are more apt to censorship the contributions by blocking accounts, deleting comments, or closing the comment section. Particularly, one of my informants experienced this censorship first hand when his account was blocked from accessing a celebrity account after he left a contradictory comment on a post. This shows that, while Twitter and Instagram livestreams working together construct the sum of one's digital identity, they are not a direct reflection of a person. It is, however, a strategically edited version of a person, specifically designed for an audience.

Audiences are more engaged with the identity-rich posts, controlled for the platforms. That is, they are more likely to engage, as measured by my engagement scores calculation (Equation 4.3 and 4.4), with the posts categorized as rich in identity affordance. This probably means that the richness of the identity affordance provides the public an impression of *being an insider* through the expression of what they truly are (Gamson, 1994). The richness in identity affordance also helps create *ambient awareness* or *digital intimacy* (Thompson, 2008). That is, it simulates a sense of being there with the posters through the little things they do. Thompson notes in his study that each post might be insignificant on its own but cumulatively creates a sophisticated portrait of the posters, "thousands of dots making a pointillist painting" (Thompson, 2008, p. 3).



### 8.3.2 Fans Interaction

As evidenced by my analyses, interactivity is lower on Instagram. The literature suggests that the platform itself is quite limited in interaction with only a few functionalities for interactions (Marwick, 2015a). In fact, the platform has an option for the account owner to view only responses from people they know. As such, it could be said that the platform is intentionally assembled for parasocial relationships – or one-sided relationships, made up of the persona who is completely unaware of the other's existence (Horton & Wohl, 1956). This is also true for Twitter, which does not employ a reciprocal-based relationship model – for example, user A can follow user B, but B does not have to follow A back. Twitter, however, provides a more interactive environment and probably simulates a sense of two-sided relationship where celebrities have more options to publicly interact with fans. For example, all the @replies and @mentions can be instantly seen on a Twitter account's homepage. This is not the case for Instagram whose algorithm organizes the comments section to show the most recent comments and/or comments from someone you are also following.

Although Twitter does not employ a reciprocal relationship, boyd et al. (2010) note that Twitter is presumed to be reciprocal in a sense that it is typical, and even expected, to @reply or @retweet someone even though you are not following them. My informants explained they expected nothing back from engaging with or following a celebrity, but they did admit that it would be nice to get a response back every once in a while. In fact, that would raise their status in a fan community (Bennett, 2014; Pegoraro, 2010). Interestingly, this implies that the interaction affordance that social media offer celebrities, subsequently offers another affordance to the fans. Although an analysis of social media affordances for fandoms would be interesting, I note this is out of my scope. This dissertation primarily looks at how social media affords

celebrities opportunities to engage in the microcelebrity practices. An example of scholarship concerning social media affordances particularly from the perspective of fandom is Ellcessor (2012) whose work suggests that the key social media affordances include the illusions of *quotidian rhythms* of interaction.

Looking at the artifacts associated with the interaction affordance, Marwick notes in her study that the uses of @mentions and @replies demonstrate the connectedness between users (Marwick, 2013). For these fans, the interaction affordance seems to afford a sense of close connection and importance, despite being one amongst millions of followers (Bennett, 2014; Hambrick, Simmons, Greenhalgh, & Greenwell, 2010; Kassing & Sanderson, 2010). My informants also reported that they would be more encouraged to engage if a celebrity had previously interacted with the public because they might be the chosen ones, too. This illustrates that when the interactions take place, they are publicly available for the observers – or a third person who is not in a conversation, but only observes the conversation (Goffman, 1959). That is, when celebrities interact with the public, it becomes part of the livestreaming, constructing one's digital presence available for the public. I also suggest that this finding is a great example of Gibson's *behavior affords behavior* phenomenon (2014a). People naturally create a mapping of actions and results by picking up information from other users. In this case, other fans learn from those with whom a celebrity interacted, and hope they would get the same results.

Again, this implies that when the interaction affordance is utilized by a celebrity, it subsequently affords another affordance, which can only be fully understood from the perspective of fandoms. I note that this is beyond the scope of my study but opens a new direction for future studies.

Interactions are probably more important for Internet celebrities who are expected to work harder to maintain the audience. Senft (2008) found in her study on *Camgirls* (i.e., female personalities who broadcasted themselves on the Web) that they often described their viewers as family and reported they felt they owed to the viewers for making them popular. On this basis, my finding indicates that Internet celebrities had more interaction-rich posts than their counterparts, although the difference was only detected on Instagram.

Regarding the consistency of the interaction practice, I found the average engagement score and changes in followers, – the two proxies of celebrity status, – were negatively related to the variance of richness scores. My informants explained that it was understandable if a celebrity did not want to interact with the public; they, however, expected the celebrities be consistent about the ways they interacted with fans. Specifically, they expected celebrities to maintain their interaction level with fans. This finding is supported by the literature suggesting that the public expects their celebrities be consistent online (Marshall, 2006; Turner, 2013). An implication from this finding is, the interaction affordance would be best utilized through a consistent appropriation.

### *8.3.3 Visibility Promotion*

The other dimension of the practice is promoting visibility beyond the existing fan base. Social media are arenas of public attention, but attention itself is a scarce resource as it gets distributed and draws on various competing issues. Certain pieces of information have to compete with others to become visible (Brighenti, 2010). Social media provide some

mechanisms that afford promoting visibility, which celebrities can appropriate to promote public exposure and grow an audience (Turner, 2010).

The analyses show that the audiences are less engaged with the posts categorized as rich in visibility promotion. My informants explained that it was understandable to see celebrities engaging in visibility-promotion activities, but could be annoying when overused. Interestingly, visibility promotion has been least mentioned by the literature. In her study, Senft explains that some viewers might profess to hate the Camgirls, seeing them as contrived and seeking attention (Senft, 2008). Although attention is essential, celebrities need to strategically engage in such activities. This is what Schwarz (2010) refers to as a *deny-and-conceal* strategy where users conceal their conscious attempts to gain followers. On this basis, Marwick documents the very same practice of an Instagram famous person who portrays herself as an ordinary girl and notes that what makes her popular might be “this seemingly effortless cool” (Marwick, 2015a, p. 150).

Further analysis on the score variance shows that the changes in followers is positively related to the variance of visibility scores. It implies that a celebrity tends to gain more followers when the visibility-rich posts are alternated. Together, this suggests that although the richness in visibility affordance might help celebrities expand their audience, it could negatively affect the existing fan base, and that celebrities should be strategic in appropriating the visibility affordance.

Hashtags, one of the artifacts associated with my visibility affordance, are frequently mentioned as a way to boost the visibility of the posts beyond the existing followers (Christensen, 2013; Page, 2012). My informants, however, reported that hashtags could be annoying. In particular, many users tend to use the popular hashtags (e.g., #followforfollow,

#love, #sun) as a way to gain likes and followers (Titlow, 2012). However, when the hashtags are irrelevant to the posts, they obviously show a conscious attempt to self-promote. This is also documented in Marwick's (2015a) study that #followforfollow or *follow for follow* has been used in more than 24 million posts and some users explicitly seek for followers by including more than 10 hashtags in their posts.

However, my informants noticed a work-around solution on Instagram through an inclusion of hashtags in the comment section. Specifically, some celebrities put the hashtags, as many as they like, as a comment rather than including them in the post. With this, the audience would not see the dozens of hashtags unless they expand and read through the comment section. Note that comments are beyond the scope of my work. Such analysis, however, would be an interesting future study as it could provide an insight into celebrity-fans relationship and social media affordances from the fans' perspective.

Interestingly, this work-around solution is a great example of Markus and Silver's (2008) argument that users do not necessarily use the artifacts as they are designed, or intended by the designer. Rather, new practices often emerge after user engagement (O'Riordan et al., 2012). In this case, the new practice is how celebrities opt to include hashtags in a comment rather than the post to appropriate the visibility affordance while concealing the attempt from the public.

#### **8.4 Microcelebrity as a Learned technique**

In the age of social media, everyone with Internet access can engage in the practice of microcelebrity but not everyone will be successful (Gamson, 1994). That is, practicing celebrity and having celebrity status are two very different things. As Marwick and boyd (2011) noted,

microcelebrity practices are *learned techniques*; they can be learned and practiced. While the literature on celebrity studies had suggested that celebrity management is a highly controlled and regulated institutional model (Marwick & boyd, 2011; Turner, 2013), social media have complicated this dynamic by creating a new set of expectations – for example, audiences may expect a more intimate relationship from a celebrity, or expect them to be more interactive on social media. Hence, it is essential that microcelebrity practitioners learn and practice the skills to appropriate social media affordances to achieve the goal of maintaining celebrity status.

In this modern era where the ability to publish searchable and enduring content has been dramatically expanded (Draper, 2016), it is important for people who are seeking for, or maintaining, their status to learn the techniques of microcelebrity. The implications from this study would be particularly useful as recommendations on best practices for anybody seeking for attention online or trying to maintain their status. Examples of the implications include an expectation of the audience to see celebrities maintain their interactional level but to strategically engage in self-promotion activities.

Nonetheless, it should be noted that adopting these best practices do not necessary guarantee the outcomes. Similar to other bottom-up or crowd-driven events, such as viral information events (Nahon & Hemsley, 2013), gaining and maintaining celebrity status are not easily controlled or predicted; by their very nature there exists a high degree of fuzziness (Nahon & Hemsley, 2013). As Collins notes, celebrity status is simply “a temporally dispensable cultural commodity” (Collins, 2008, p. 102).

An alternative explanation is the concept of *the-rich-get-richer*, meaning it is usually easier for those at the top (highly followed users) to expand the network than those who are less

followed (Barabasi, 2003). By the very concept, it also raises a barrier for *newcomers*. This essentially leads the numbers of followers of most networks to exhibit a power-law distribution where there is a relatively much smaller number of highly followed users. Even if newcomers pass the barrier, i.e., successfully gain a substantial audience, they are subjected to the *fifteen-minutes-of-fame*, a short-lived celebrity status. The question remains: can they maintain the status? The maintenance of the status indeed requires *labor* from the celebrities themselves (Mavroudis & Milne, 2016) as well as the public, consciously or not (Abidin, 2016) all of which is mediated by the affordances of platforms like Twitter and Instagram.

## 8.5 Limitations and Future work

This study has a few limitations. First, the performance of celebrity may be different on different platforms and this study only examined the practices on Twitter and Instagram. As such, I do not claim that my findings can be generalized beyond Twitter and Instagram. However, future studies can adopt a similar methodological model and make use of my framework, but compare activities on different platforms – for example, Facebook and YouTube which are also interesting venues for people seeking for audience and celebrity status.

Second, majority of the AMT workers who annotated the richness labels identified themselves as not frequently posting on social media. Specifically, 23.43% of the Twitter labelling workers categorized themselves as a frequent poster and it was 34.35% for the Instagram labelling task. I note that their nature as a passive user could potentially limit their ability to justify the richness of the posts. However, previous studies show that majority of social media population only passively consumes rather than creating content (Brandtzæg, 2012;

Reuter, Heger, & Pipek, 2013). As such, this group of AMT workers is a good representative of the actual social media population, including celebrities' audiences, in terms of social media usage.

Third, my richness score analyses were conducted on the celebrities' posts only. The data did not include comments associated with the posts. Analyzing comments from the public could be an interesting piece of analysis. Such work is possible with the use of the framework and would be an interesting future study as it could provide an insight into celebrity-fan relationships and social media affordances from the fans' perspective.

It should be noted that Twitter had the 140-characters limit as of the time of the data collection. It was changed to 280-characters in November 2017. I argue that the change does not impact the analyses and results of my study, but demonstrates the flexibility of the framework. Specifically, the theoretical foundation of Affordances makes the framework independent of the technological features of the sites. This is essential as the environments of social media have been rapidly evolving with the progressive development of new features (Bruns & Burgess, 2011).

In November 2016, Instagram launched a new feature to let users share *Stories* – or posts that would last only 24 hours and disappear. Although they would be useful for my analysis, they were not included in my data collection. Stories were intentionally designed to be ephemeral and *off the record*. Even from the Instagram users' point of view, Stories can be replayed only once and they will disappear.

Lastly, I did not control for the number of followers a celebrity has. This reflects a conscious choice. By adopting the perspective of celebrity as practice (Marwick & boyd, 2011),



we can eschew a process of selecting actors with more than some arbitrary number of followers. Rather, different celebrity types will operate in larger and smaller environments, which may be reflected in how they emphasize one affordance dimension over another. Although one might expect that Internet celebrities would operate in a smaller environment as they are less famous than their counterparts, the interviews with audience members show that the public did not see much difference between the celebrity types. This, in fact, echoes Marwick and boyd's (2011) perspective of celebrity as a continuum between globally famous down to a local/niche celebrity, rather than a binary quality (i.e., you are or you are not a celebrity).

## CHAPTER 9

### CONCLUSION

In this study, I examined microcelebrity practices on multiple social media through the theoretical lens of Affordances. The theory of microcelebrity explains how ordinary people turn their public persona into media content to be consumed by an audience with the goal of gaining and/or maintaining their audience, who are regarded as fans. To accomplish this, the theory suggests that people employ a set of online self-presentation techniques that typically consist of three core practices: identity constructions, fan interactions, and visibility promotion. Studies on single platforms (e.g., Twitter), however, show that not all microcelebrities necessarily engage in all core practices to the same degree. Importantly, celebrities are increasingly using multiple platforms simultaneously to expand their audience while overcoming the limitations of a particular platform. This points to a gap in the literature and calls for a cross-platform study.

This dissertation employs a mixed-methods research design to reveal how social media platforms, i.e., Twitter and Instagram, afford mechanisms for celebrities to grow and maintain their audience by strategically utilizing identity, interaction, and visibility affordance. The first phase of the study relies on a richness framework that quantifies social media activities to measurable richness constructs. The framework was developed through a conceptual lens of Affordance theory (Gibson, 2014a) and borrowing a notion from Media Richness theory (Daft & Lengel, 1986).

Specifically, I suggest that although the theory of Affordances provides a useful foundation for assessing information environments from the action possibilities perspective, it does not provide a systematic way to examine how the affordances, when undertaken, enable users to engage in social activities. As such, I developed the notion of *affordance richness* as the ability of a medium (i.e., tweet or Instagram post) to deliver the information necessary in affording a particular action by using artifacts of social media. Just like media richness is the ability of a medium to deliver rich information (Daft & Lengel, 1986), affordance richness measures the richness of affordances made possible by a medium like tweets and Instagram posts.

With this framework, I generated the richness scores of a large set of Twitter and Instagram data from celebrities of both types: mainstream and Internet famous, and performed a series of quantitative analyses on the richness scores. Each of the analyses was designed to address different research questions regarding social media usage by different groups of celebrities and how audience responded to different microcelebrity strategies.

Specifically, RQ1 asked if the practices of microcelebrity were different by platforms (i.e., Twitter vs. Instagram). A series of paired *t*-tests show that the practices differ: celebrities tend to create more identity-rich and visibility-rich posts on Twitter but more interaction-rich posts on Instagram. Next, RQ2 asked if the practices were different by celebrity types (i.e., mainstream vs. Internet famous). Another series of *t*-tests indicate that the only significant difference is the interaction practice on Instagram, where Internet celebrities tend to create interaction-rich posts more often than their counterparts. RQ3 asked if the audiences responded to the two types of celebrity similarly or differently. Specifically, I operationalized *audience responses* as the engagement score and changes in followers and constructed linear regression

models. The results indicate that audiences tend to be more engaged with identity-rich and interaction-rich posts but less engaged with visibility-rich posts. However, the richness scores have no significant effects on changes in followers. Next, RQ4 asked how the audiences responded to changes in microcelebrity strategies. For this question, I operationalized the changes in strategies as variance scores. The linear regression models indicate that the mean engagement score and changes in followers are negatively related to the variance of the interaction scores. However, the variance of visibility scores is positively related to the changes in followers.

The aforementioned findings informed the design of the follow-up interviews with audience members or fans. This is particularly important because understanding expectations and behaviors of fans is relevant not only as a means to enhance the practice outcome and sustain promotional activity (Usher, 2015), but also because it contributes to our understandings about contemporary celebrity-fan relationships mediated by social media. The qualitative phase also helped answer the last question regarding audience expectations and the reasons behind their engagements with the celebrities.

The interview protocol consisted of eight main topics such as the uses of social media, perceptions about celebrity and perceptions of the core practices on microcelebrities. In total, I conducted 15 one-on-one interviews with audience members, each of which was roughly an hour long. Coding of the interview transcripts resulted in three main themes, each of which consists of six codes. The findings from this qualitative study were used as a supplement to the findings of the prior quantitative analyses. For example, from the quantitative analyses I found that the engagement score was negatively related to the variance of interaction richness scores. This was

understandable when the informants explained that they expected to see celebrities maintain a steady interactional level with fans.

This study makes contributions to the theory of Microcelebrity and offers practical contributions by providing broad insights from both practitioners' and audiences' perspectives. This is essentially important given that microcelebrity is a learned practice rather than an inborn trait. The study also makes a methodological contribution through the development of its framework. Specifically, this framework can be a tool to study social media activities from the perspective of technological affordances and provides a way for researchers to compare actors in different contexts. Lastly, this work also makes theoretical contributions to a growing body of literature around the theory of Affordances with the development of the notion of *affordance richness*. Researchers can leverage the notion to explain the ability of an environment to afford users to perform an activity along multiple dimensions of the practices. This could be particularly useful when the object can be used in different ways, or when the phenomenon of interest contains multiple dimensions of practices. Examples of such phenomena include microcelebrity, composed of three core practices.

## APPENDICES

### Recruiting Letter

#### **Microcelebrity Practices: Towards Cross-Platform Studies**

#### **Recruiting Letter**

Jeff Hemsley; Sikana Tanupabrungsun

Dear [NAME]

My name is Sikana Tanupabrungsun and I am working with Dr. Jeff Hemsley, who is a Professor at Syracuse University. We are currently working on a research project to understand fan-celebrity relationship on social media like Twitter and Instagram.

As such, we would give you a \$15 Amazon gift card if you would be willing to participate in a roughly 50-60 minutes audio interview with us. Please be informed that the interview will be audio recorded.

To schedule an interview with us, please email me at [stanupab@syr.edu](mailto:stanupab@syr.edu) and we can send you more details.

Thank you!

All the best,

Sikana Tanupabrungsun

Ph.D. Student

School of Information Studies

Syracuse University

[stanupab@syr.edu](mailto:stanupab@syr.edu)

## Interview Protocol

Name:

Date:

### Demographic

1. How old are you?
2. What is your gender?
3. What is your educational background?
4. Tell us more about yourself. We would like to know more about you, your life experiences, what you have done, where you work, and more.

### General information about social media uses

1. What social media are you currently using?
2. Which ones do you prefer the most?
3. What are the main reasons you use social media?
4. Do you have accounts on Twitter and/or Instagram? If both, which one do you like better? If not, why do you prefer one platform over the other?
5. How many followers do you have on Twitter and/or Instagram?
6. How many accounts are you following on Twitter and/or Instagram?
7. How long have you been on Twitter/Instagram?
8. Why did you decide to join in the beginning?
9. Could you please describe your uses of Twitter/Instagram?
10. Would you categorize yourself as a poster, liker or lurker?
11. In what ways do you use Twitter/Instagram the most?

Perception of celebrity

1. In your own words, what does 'celebrity' mean?
2. In your opinion, what are the differences between mainstream and Internet celebrities?

Use of social media in the celebrity context

1. What percent of accounts that you follow on Twitter/Instagram would you classify as celebrities or famous people?
2. How many of them are Internet celebrities? Can you give me some examples of such accounts?
3. How many of them are mainstream celebrities? Can you give me some examples of such accounts?
4. Why do you follow celebrities? What do you expect from following them?
5. How do your expectations differ by celebrity types i.e., mainstream and Internet celebrity?
6. Do you follow the same persons on both platforms? Why/why not?
7. What are the reasons you follow some celebrities but not others?
8. What are the reasons you unfollow someone?
9. Why would you 'like' or 'retweet' some posts but not others?
10. Why would you 'reply' to or 'comment' on some posts but not others?
11. Are there any limitations of Twitter/Instagram that prevent some activities of celebrity-fan relationships?
12. Are there anything you want to change about the platforms?

Perception of identity work

1. In what ways do you see celebrities present their personality or character?



2. How do they express themselves?
3. Are there any differences between celebrity types i.e., mainstream and Internet famous?
4. When celebrities do *X* and *Y* [*using the answers from previous question*], how do they help you imagine or create a picture of their character or personality?
5. How do *X* and *Y* [*using the answers from the first question*] affect your decision to like, reply, follow or unfollow? Are the effects different by celebrity types i.e., mainstream and Internet famous?
6. Are they consistently doing *X* and *Y* [*using the answers from the first question*] over the course of time you are following them? Are they consistent across platforms?
7. Would you unfollow them if they stop doing *X* and *Y* [*using the answers from previous question*]?
8. What are some features of Twitter and Instagram you think are helpful for expressing one's personality/character?

#### Perception of interaction work

1. Do you see celebrities frequently interacting with others? In what ways do they interact with others?
2. Do they interact with fans or just other famous people? How do you feel about that?
3. Would you feel differently if they did it the other way around?
4. How do the interactions affect your decision to like, reply, follow or unfollow? Are the effects different by celebrity types i.e., mainstream and Internet famous?
5. Are they consistently doing *X* and *Y* [*using the answers from the first question*] over the course of time you are following them? Are they consistent across platforms? How does that affect your decision to like, reply, follow or unfollow the accounts?

6. Have you attempted to get responses from them? How?
7. Would you unfollow them if they stop interacting with fans?
8. What are some features of Twitter and Instagram you think are helpful for celebrities interacting with fans?

#### Perception of visibility work

1. How did you find or get to know these celebrity accounts?
2. What did they do to promote themselves?
3. When they do *X* and *Y* [*using the answers from previous question*] affect your decision to like, reply, follow or unfollow? Are the effects different by celebrity types?
4. Are they consistently doing *X* and *Y* [*using the answers from the second question*] over the course of time you are following them? Are they consistent across platforms? How does that affect your decision to like, reply, follow or unfollow the accounts?
5. What else they could do to become even more famous?
6. What are some features of Twitter and Instagram you think are helpful for celebrities promoting their accounts beyond the existing fan base?

#### Fan-celebrity community

1. What do the fan communities look like?
2. Do you feel you are part of the community/communities?
3. What actions do you typically take to interact with other fans?
4. Compare to the pre-social media era, what have been changed in terms of fan-celebrity interactions? What have not?

## Codebook

Codes	Definition	Example
<i>1. Microcelebrity practices on different platforms and of different celebrity types</i>		
1.1 Presenting persona/identity on Twitter vs. Instagram	This code explains if one platform is more appropriate than the other for celebrities to undertake the identity affordance.	<i>I feel like it is tougher to do it on Twitter. I know you can share your pictures on Twitter but I know it's mostly word based.</i>
1.2 Interacting with fans on Twitter vs. Instagram	This code explains if one platform is more appropriate than the other for celebrities to undertake the interaction affordance.	<i>I also happen to see this [interaction] happens on TW more often than other platforms.</i>
1.3 Promoting visibility on Twitter vs. Instagram	This code explains if one platform is more appropriate than the other for celebrities to undertake the visibility affordance.	<i>Twitter, I think they can promote their own post or they can tweet in a high frequency that they will always be in my 'tweets you are missing' and that's usually the case for most people.</i>

1.4 Internet vs. Mainstream celebrities presenting persona/identity	This code explains if different types of celebrity appropriate the identity affordance similarly or differently.	<i>I'd suspect Internet celebrities are probably more involved presenting themselves on Instagram and social media versus non-Internet famous.</i>
1.5 Internet vs. Mainstream celebrities interacting with fans	This code explains if different types of celebrity appropriate the interaction affordance similarly or differently.	<i>I'd definitely say Internet celebs comment or responds to the fans more than regular celeb.</i>
1.6 Internet vs. Mainstream celebrities promoting visibility	This code explains if different types of celebrity appropriate the visibility affordance similarly or differently.	<i>I mean I think the Internet celeb, my expectations are they are going to be more active, like trying harder to get followers.</i>
<b><i>2. The effects of microcelebrity strategies on maintaining and growing an audience</i></b>		
2.1 Effects of presenting persona/identity on engagement decisions	This code explains the effects of the identity-construction practices on engagement decisions.	<i>When I see a picture of Justin Bieber, I'd go crazy. And I like to click like because of a particular person that is on there.</i>

<p>2.2 Effects of interacting with fans on engagement decisions</p>	<p>This code explains the effects of the interactional practices on engagement decisions.</p>	<p><i>I guess if I felt the need to interact with the post like if I have a comment, I might then go back and do something and check like if he responses to a lot of posts, I feel ok about commenting on it too.</i></p>
<p>2.3 Effects of promoting visibility on engagement decisions</p>	<p>This code explains the effects of the visibility-promotion practices on engagement decisions.</p>	<p><i>If I think there are too many unnecessary hashtags, there's a slight chance that I'll just skip it just because it is kind of annoying.</i></p>
<p>2.4 Effects of presenting persona/identity on follow and unfollow decisions</p>	<p>This code explains the effects of the identity-construction practices on follow/unfollow decisions.</p>	<p><i>I don't expect, they don't owe me anything, they don't owe me any sort of life updates.</i></p>
<p>2.5 Effects of interacting with fans on follow and unfollow decisions</p>	<p>This code explains the effects of the interactional practices on follow/unfollow decisions.</p>	<p><i>I don't care who they are talking to really. I don't, they won't make me unfollow them.</i></p>

2.6 Effects of promoting visibility on follow and unfollow decisions	This code explains the effects of the visibility-promotion practices on follow/unfollow decisions.	<i>I mean yea I sometimes got to know the accounts by exploring the hashtags like when they used the clothing brand hashtags or something that I like.</i>
<b>3. The Effects of the consistency in microcelebrity strategies on maintaining and growing an audience</b>		
3.1 Effects of the consistency in presenting persona/identity on engagement decisions	This code explains the effects of the consistency in identity-construction practices on engagement decisions.	NA
3.2 Effects of the consistency in interacting with fans on engagement decisions	This code explains the effects of the consistency in interactional practices on engagement decisions.	<i>The more I see somebody interacting with people that are commenting, the more I want to comment just because I know that it [fan interaction] is something they truly do care with and that they like to respond.</i>

<p>3.3 Effects of the consistency in promoting visibility on engagement decisions</p>	<p>This code explains the effects of the consistency in visibility-promotion practices on engagement decisions.</p>	<p><i>It's definitely annoying when celeb does that [self-promotion] a lot and yea, sometimes I just navigate away like skip it.</i></p>
<p>3.4 Effects of the consistency in presenting persona/identity on follow and unfollow decisions</p>	<p>This code explains the effects of the consistency in identity-construction practices on follow/unfollow decisions.</p>	<p><i>NA</i></p>
<p>3.5 Effects of the consistency in interacting with fans on follow and unfollow decisions</p>	<p>This code explains the effects of the consistency in interactional practices on follow/unfollow decisions.</p>	<p><i>If they never interact with people before, I'll be like fine that person doesn't really interact. But if they used to and now they don't, I would unfollow them.</i></p>
<p>3.6 Effects of the consistency in promoting visibility on follow and unfollow decisions</p>	<p>This code explains the effects of the consistency in visibility-promotion practices on follow/unfollow decisions.</p>	<p><i>I sometimes unfollow people who crave attention like all the time, this is especially true for celebs, I think. It's kinda annoying.</i></p>

# IRB Exemption

**SYRACUSE UNIVERSITY**



**INSTITUTIONAL REVIEW BOARD  
MEMORANDUM**

**TO:** Jeff Hemsley  
**DATE:** October 11, 2017  
**SUBJECT:** Determination of Exemption from Regulations  
**IRB #:** 17-323  
**TITLE:** *Microcelebrity Practices: Towards Cross-Platform Studies*

The above referenced application, submitted for consideration as exempt from federal regulations as defined in 45 C.F.R. 46, has been evaluated by the Institutional Review Board (IRB) for the following:

1. determination that it falls within the one or more of the five exempt categories allowed by the organization;
2. determination that the research meets the organization's ethical standards.

It has been determined by the IRB this protocol qualifies for exemption and has been assigned to category **2**. This authorization will remain active for a period of five years from **October 6, 2017** until **October 5, 2022**.

**CHANGES TO PROTOCOL:** Proposed changes to this protocol during the period for which IRB authorization has already been given, cannot be initiated without additional IRB review. If there is a change in your research, you should notify the IRB immediately to determine whether your research protocol continues to qualify for exemption or if submission of an expedited or full board IRB protocol is required. Information about the University's human participants protection program can be found at: <http://orip.syr.edu/human-research/human-research-irb.html> Protocol changes are requested on an amendment application available on the IRB web site; please reference your IRB number and attach any documents that are being amended.

**STUDY COMPLETION:** Study completion is when all research activities are complete or when a study is closed to enrollment and only data analysis remains on data that have been de-identified. A Study Closure Form should be completed and submitted to the IRB for review ([Study Closure Form](#)).

Thank you for your cooperation in our shared efforts to assure that the rights and welfare of people participating in research are protected.

Tracy Cromp, M.S.W.  
Director

**DEPT:** Information Studies, 346 Hinds Hall

**STUDENT:** Sikana Tanupabrungsun



## REFERENCES

- Abidin, C. (2014). #In\$tagLam: Instagram as a Repository of Taste, a Burgeoning Marketplace, a War of Eyeballs. In M. Berry & M. Schleser (Eds.), *Mobile Media Making in an Age of Smartphones* (pp. 119–128). Palgrave Macmillan US.  
[https://doi.org/10.1057/9781137469816\\_11](https://doi.org/10.1057/9781137469816_11)
- Abidin, C. (2015). Micro-microcelebrity: Branding Babies on the Internet. *M/C Journal*, 18(5). Retrieved from <http://journal.media-culture.org.au/index.php/mcjournal/article/view/1022>
- Abidin, C. (2016). Visibility labour: Engaging with Influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100.  
<https://doi.org/10.1177/1329878X16665177>
- Ahmed, A. (2012). Hypothesizing the aptness of social media and the information richness requirements of disaster management. *ECIS 2012 Proceedings*. Retrieved from <http://aisel.aisnet.org/ecis2012/157>
- Araújo, C. S., Corrêa, L. P. D., da Silva, A. P. C., Prates, R. O., & Meira, W. (2014). It is not just a picture: revealing some user practices in instagram (pp. 19–23). Presented at the Web Congress (LA-WEB), 2014 9th Latin American, IEEE.
- Arguello, J., Butler, B. S., Joyce, E., Kraut, R., Ling, K. S., Rosé, C., & Wang, X. (2006). Talk to me: foundations for successful individual-group interactions in online communities (pp. 959–968). Presented at the Proceedings of the SIGCHI conference on Human Factors in computing systems, ACM.
- Ausserhofer, J., & Maireder, A. (2013). National Politics on Twitter. *Information, Communication & Society*, 16(3), 291–314.  
<https://doi.org/10.1080/1369118X.2012.756050>

- Bakhshi, S., Shamma, D. A., & Gilbert, E. (2014). Faces Engage Us: Photos with Faces Attract More Likes and Comments on Instagram. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (pp. 965–974). New York, NY, USA: ACM. <https://doi.org/10.1145/2556288.2557403>
- Barabasi, A.-L. (2003). *Linked: How Everything Is Connected to Everything Else and What It Means*. Plume. Retrieved from <http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0452284392>
- Bennett, L. (2014). Fan/celebrity interactions and social media: connectivity and engagement in Lady Gaga fandom. *The Ashgate Research Companion to Fan Cultures, Surrey: Ashgate Publishing, Ltd*, 109–120.
- Bennett, L., Segerberg, A., & Walker, S. (2014). Organization in the crowd: peer production in large-scale networked protests. *Information, Communication & Society*, *17*(2), 232–260. <https://doi.org/10.1080/1369118X.2013.870379>
- Black, A., Mascaro, C., Gallagher, M., & Goggins, S. P. (2012). Twitter zombie: Architecture for capturing, socially transforming and analyzing the Twittersphere (pp. 229–238). Presented at the Proceedings of the 17th ACM international conference on Supporting group work, ACM.
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310.
- Bloomfield, B. P., Latham, Y., & Vurdubakis, T. (2010). Bodies, technologies and action possibilities: When is an affordance? *Sociology*, *44*(3), 415–433.

- boyd, danah. (2008). *Taken out of context: American teen sociality in networked publics* (PhD Dissertation). University of California-Berkeley, School of Information, Berkeley, CA.
- boyd, danah. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In *A Networked self: identity, community, and culture on social network sites* (pp. 39–58). New York: Routledge.
- boyd, danah, Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *System Sciences (HICSS), 2010 43rd Hawaii International Conference on* (pp. 1–10). IEEE Computer Society.
- Bradner, E. (2001). Social affordances of computer-mediated communication technology: understanding adoption (pp. 67–68). Presented at the CHI'01 Extended abstracts on human factors in computing systems, ACM.
- Bradner, E., Kellogg, W. A., & Erickson, T. (1999). Social affordances of Babble: A field study of chat in the workplace (Vol. 99, pp. 12–16). Presented at the ECSCW.
- Brandtzæg, P. B. (2012). Social Networking Sites: Their Users and Social Implications — A Longitudinal Study. *Journal of Computer-Mediated Communication*, 17(4), 467–488. <https://doi.org/10.1111/j.1083-6101.2012.01580.x>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140.
- Brighenti, A. M. (2010). *Visibility in social theory and social research*. Springer.
- Brinker, D. L., Gastil, J., & Richards, R. C. (2015). Inspiring and Informing Citizens Online: A Media Richness Analysis of Varied Civic Education Modalities. *Journal of Computer-Mediated Communication*, 20(5), 504–519. <https://doi.org/10.1111/jcc4.12128>

- Bruns, A., & Burgess, J. E. (2011). The use of Twitter hashtags in the formation of ad hoc publics. In *Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011*.
- Bucher, T., & Helmond, A. (2017). The affordances of social media platforms. *SAGE Handbook of Social Media*. London: Sage.
- Burgess, J., & Green, J. (2009). *YouTube: Online Video and Participatory Culture* (1st ed.). Cambridge: Polity.
- Burgess, J., Mitchell, P., & Münch, F. (2018). Social media rituals: The uses of celebrity death in digital culture. In *A Networked Self: Birth, Life, Death*. Taylor & Francis Group, LLC.
- Burke, M., & Kraut, R. (2008). Mind your Ps and Qs: the impact of politeness and rudeness in online communities (pp. 281–284). Presented at the Proceedings of the 2008 ACM conference on Computer supported cooperative work, ACM.
- Carlson, J. R., & Zmud, R. W. (1999). Channel Expansion Theory and the Experiential Nature of Media Richness Perceptions. *Academy of Management Journal*, 42(2), 153–170.  
<https://doi.org/10.2307/257090>
- Centeno, D. D. G. (2010). Celebri-fication in Philippine Politics: Exploring the Relationship Between Celebrity Endorser's Parasociability and the Public's Voting Behavior. *Social Science Diliman*, 6(1).
- Christensen, C. (2013). Wave-Riding and Hashtag-Jumping. *Information, Communication & Society*, 16(5), 646–666. <https://doi.org/10.1080/1369118X.2013.783609>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- Collins, S. (2008). Making the Most out of 15 Minutes: Reality TV's Dispensable Celebrity. *Television & New Media*, 9(2), 87–110. <https://doi.org/10.1177/1527476407313814>

- Conway, B. A., Kenski, K., & Wang, D. (2013). Twitter Use by Presidential Primary Candidates During the 2012 Campaign. *American Behavioral Scientist*, 57(11), 1596–1610.  
<https://doi.org/10.1177/0002764213489014>
- Cottom, T. M. (2015). “Who The Fuck Do You Think You Are?” Academic Engagement, Microcelebrity and Digital Sociology from the Far Left of the Matrix of Domination. *Ada: A Journal of Gender, New Media, and Technology*.  
<https://doi.org/10.7264/N3319T5T>
- Coyle, J. R., & Thorson, E. (2001). The effects of progressive levels of interactivity and vividness in Web marketing sites. *Journal of Advertising*, 30(3), 65–77.
- Creswell, J. W. (1994). Research design: Quantitative and qualitative approaches. *Thousand Oakes: Sage Publication*.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.
- Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, 32(5), 554–571.  
<https://doi.org/10.1287/mnsc.32.5.554>
- Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message Equivocality, Media Selection, and Manager Performance: Implications for Information Systems. *MIS Quarterly*, 11(3), 355–366. <https://doi.org/10.2307/248682>
- Dao, D. V. (2015). Social Media Classification Scheme in Online Teaching and Learning Activities: A Consideration for Educators. *International Journal of Education and Social Science*, 2(4), 85–94.

- Darling, E., Shiffman, D., Côté, I., & Drew, J. (2013). The role of Twitter in the life cycle of a scientific publication. *Ideas in Ecology and Evolution*, 6(1). Retrieved from <http://ojs.library.queensu.ca/index.php/IEE/article/view/4625>
- Dennis, A. R., & Kinney, S. T. (1998). Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information Systems Research*, 9(3), 256–274.
- Dennis, A. R., & Kinney, S. T. (1998). Testing Media Richness Theory in the New Media: The Effects of Cues, Feedback, and Task Equivocality. *Information Systems Research*, 9(3), 256–274. <https://doi.org/10.1287/isre.9.3.256>
- Dennis, A. R., & Valacich, J. S. (1999). Rethinking media richness: towards a theory of media synchronicity. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1999. HICSS-32* (Vol. Track1, p. 10 pp.-). <https://doi.org/10.1109/HICSS.1999.772701>
- 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. <https://doi.org/10.1016/j.chb.2016.11.009>
- Draper, N. (2016). REPUTATION, INC.: THE INDUSTRIALIZATION OF DIGITAL SELF PRESENTATION AND ONLINE PRIVACY. *AoIR Selected Papers of Internet Research*, 5.
- Driessens, O. (2013). The celebritization of society and culture: Understanding the structural dynamics of celebrity culture. *International Journal of Cultural Studies*, 16(6), 641–657.

- Du, S., & Vieira, E. T. (2012). Striving for Legitimacy Through Corporate Social Responsibility: Insights from Oil Companies. *Journal of Business Ethics, 110*(4), 413–427.  
<http://dx.doi.org.libezproxy2.syr.edu/10.1007/s10551-012-1490-4>
- Ellessor, E. (2012). Tweeting@ feliciaday: Online social media, convergence, and subcultural stardom. *Cinema Journal, 51*(2), 46–66.
- Ellessor, E. (2018). “One tweet to make so much noise”: Connected celebrity activism in the case of Marlee Matlin. *New Media & Society, 20*(1), 255–271.  
<https://doi.org/10.1177/1461444816661551>
- Ellison, N. B., & Vitak, J. (2015). Social network site affordances and their relationship to social capital processes. *The Handbook of the Psychology of Communication Technology, 32*, 205–228.
- El-Shinnawy, M. M., & Markus, M. L. (1992). Media richness theory and new electronic communication media: A study of voice mail and electronic mail (pp. 91–105). Presented at the ICIS.
- El-Shinnawy, M., & Markus, M. L. (1997). The poverty of media richness theory: explaining people’s choice of electronic mail vs. voice mail. *International Journal of Human-Computer Studies, 46*(4), 443–467. <https://doi.org/10.1006/ijhc.1996.0099>
- Enli, G. S., & Skogerbø, E. (2013). Personalized Campaigns in Party-Centred Politics. *Information, Communication & Society, 16*(5), 757–774.  
<https://doi.org/10.1080/1369118X.2013.782330>
- Estabrooks, A., Jo, T., & Japkowicz, N. (2004). A multiple resampling method for learning from imbalanced data sets. *Computational Intelligence, 20*(1), 18–36.

- Facebook. (2017). Help Center. Retrieved April 3, 2017, from <https://www.facebook.com/help/794890670645072>
- Faraway, J. J. (2004). *Linear models with R* (Vol. 63). Chapman and Hall/CRC.
- Fayard, A.-L., & Weeks, J. (2014). Affordances for practice. *Information and Organization*, 24(4), 236–249. <https://doi.org/10.1016/j.infoandorg.2014.10.001>
- Frederick, E., Lim, C. H., Clavio, G., Pedersen, P. M., & Burch, L. M. (2014). Choosing Between the One-Way or Two-Way Street: An Exploration of Relationship Promotion by Professional Athletes on Twitter. *Communication & Sport*, 2(1), 80–99. <https://doi.org/10.1177/2167479512466387>
- Freund, Y., & Schapire, R. E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory* (pp. 23–37). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/3-540-59119-2\\_166](https://doi.org/10.1007/3-540-59119-2_166)
- Fulk, J., & Boyd, B. (1991). Emerging Theories of Communication in Organizations. *Journal of Management*, 17(2), 407.
- Fulk, J., Schmitz, J., & Steinfield, C. (1990). A Social Influence Model of Technology use. In J. Fulk & C. Steinfield, *Organizations and Communication Technology* (pp. 117–140). 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. Retrieved from <http://sk.sagepub.com/books/organizations-and-communication-technology/n6.xml>
- Gamson, J. (1994). *Claims to fame: Celebrity in contemporary America*. Univ of California Press.
- Gamson, J. (2011). The unwatched life is not worth living: The elevation of the ordinary in celebrity culture. *PMLA*, 126(4).



- Gao, F. (2016). Social Media as a Communication Strategy: Content Analysis of Top Nonprofit Foundations' Micro-blogs in China. *International Journal of Strategic Communication*, 10(4), 255–271. <https://doi.org/10.1080/1553118X.2016.1196693>
- Gaver, W. W. (1991). Technology affordances (pp. 79–84). Presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, ACM.
- Gibson, J. J. (2014a). *The ecological approach to visual perception: classic edition*. Psychology Press.
- Gibson, J. J. (2014b). The Theory of Affordances. *The People, Place, and Space Reader*, 56.
- Gilbert, E., Bakhshi, S., Chang, S., & Terveen, L. (2013). I need to try this?: a statistical overview of pinterest (pp. 2427–2436). Presented at the Proceedings of the SIGCHI conference on human factors in computing systems, ACM.
- Glassman, M., Straus, J. R., & Shogan, C. J. (2009). Social networking and constituent communication: Member use of Twitter during a two-week period in the 111th Congress. Congressional Research Service, Library of Congress.
- Gleason, B. (2013). # Occupy Wall Street: Exploring informal learning about a social movement on Twitter. *American Behavioral Scientist*, 57(7), 966–982.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, N.Y.: Doubleday.
- Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, 61(8), 1612–1621. <https://doi.org/10.1002/asi.21344>
- Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198–208. <https://doi.org/10.1177/0165551506062337>

- Graham, T., Broersma, M., Hazelhoff, K., & Haar, G. van 't. (2013). Between Broadcasting Political Messages and Interacting with Voters. *Information, Communication & Society, 16*(5), 692–716. <https://doi.org/10.1080/1369118X.2013.785581>
- Grant, W. J., Moon, B., & Grant, J. B. (2010). Digital Dialogue? Australian Politicians' use of the Social Network Tool Twitter. *Australian Journal of Political Science, 45*(4), 579–604. <https://doi.org/10.1080/10361146.2010.517176>
- Graves, L. (2007). The affordances of blogging: A case study in culture and technological effects. *Journal of Communication Inquiry, 31*(4), 331–346.
- Greenwood, S., Perrin, rew, & Duggan, M. (2016, November 11). Social Media Update 2016. Retrieved March 14, 2017, from <http://www.pewinternet.org/2016/11/11/social-media-update-2016/>
- Gregory, K. (2018). Anger in Academic Twitter: Sharing, Caring, and Getting Mad Online. *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society, 16*(1), 176–193.
- Hackley, C., Hackley, R. A., & Bassiouni, D. H. (2017). Implications of the selfie for marketing management practice in the era of celebrity. *Marketing Intelligence & Planning, 36*(1), 49–62. <https://doi.org/10.1108/MIP-07-2017-0124>
- 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.
- Hearn, A. (2008). Meat, Mask, Burden Probing the contours of the brandedself. *Journal of Consumer Culture, 8*(2), 197–217.

- Hemphill, L., Otterbacher, J., & Shapiro, M. (2013). What's congress doing on twitter? In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 877–886). ACM.
- Hemsley, J. (2016). Studying the viral growth of a connective action network using information event signatures. *First Monday*, 21(8). <https://doi.org/10.5210/fm.v21i8.6650>
- Honey, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on* (pp. 1–10). IEEE.
- Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on* (pp. 1–10). IEEE.
- Hong, L., Dan, O., & Davison, B. D. (2011). Predicting popular messages in twitter (pp. 57–58). Presented at the Proceedings of the 20th international conference companion on World wide web, ACM.
- Horton, D., & Wohl, R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, 19(3), 215–229.
- Hsieh, Y. P. (2012). Online social networking skills: The social affordances approach to digital inequality. *First Monday*, 17(4).
- Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What We Instagram: A First Analysis of Instagram Photo Content and User Types. Presented at the ICWSM.
- Huang, J., Thornton, K. M., & Efthimiadis, E. N. (2010). Conversational tagging in twitter. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia* (pp. 173–178). ACM.

- Huba, J. (2013). *Monster Loyalty: How Lady Gaga Turns Followers into Fanatics*. Penguin.
- Huberman, B. A., Romero, D. M., & Wu, F. (2008). *Social Networks that Matter: Twitter Under the Microscope* (SSRN Scholarly Paper No. ID 1313405). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=1313405>
- Hutchby, I. (2014). Communicative affordances and participation frameworks in mediated interaction. *Journal of Pragmatics*, 72, 86–89.  
<https://doi.org/10.1016/j.pragma.2014.08.012>
- Hutto, C. J., Yardi, S., & Gilbert, E. (2013). A Longitudinal Study of Follow Predictors on Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 821–830). New York, NY, USA: ACM.  
<https://doi.org/10.1145/2470654.2470771>
- Jalonen, H. (2014). Social media and emotions in organisational knowledge creation. In *2014 Federated Conference on Computer Science and Information Systems (FedCSIS)* (pp. 1371–1379). <https://doi.org/10.15439/2014F39>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
- Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. NYU press.
- Jerslev, A., & Mortensen, M. (2016). What is the self in the celebrity selfie? Celebriphatic communication and performativity. *Celebrity Studies*, 7(2), 249–263.
- Jung, S. (2011). K-pop, Indonesian fandom, and social media. *Transformative Works and Cultures*, 8(1). <https://doi.org/10.3983/twc.2011.0289>

- Jung, Y., Tay, A., Hong, T., Ho, J., & Goh, Y. H. (2017). Politician's Strategic Impression Management on Instagram. Presented at the Proceedings of the 50th Hawaii International Conference on System Sciences.
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59–68.  
<https://doi.org/10.1016/j.bushor.2009.09.003>
- Kaplan, A. M., & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, 54(2), 105–113.  
<https://doi.org/10.1016/j.bushor.2010.09.004>
- Kaptelinin, V., & Nardi, B. (2012). Affordances in HCI: toward a mediated action perspective. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 967–976). ACM.
- Kassing, J. W., & Sanderson, J. (2010). Fan-athlete interaction and Twitter tweeting through the Giro: A case study. *International Journal of Sport Communication*, 3(1), 113–128.
- Khamis, S., Ang, L., & Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of Social Media Influencers. *Celebrity Studies*, 8(2), 191–208.
- Kinney, S. T., & Watson, R. T. (1992). The effect of medium and task on dyadic communication.
- Kishi, M. (2008). Perceptions and use of electronic media: Testing the relationship between organizational interpretation differences and media richness. *Information & Management*, 45(5), 281–287. <https://doi.org/10.1016/j.im.2008.02.008>
- Kivran-Swaine, F., Govindan, P., & Naaman, M. (2011). The Impact of Network Structure on Breaking Ties in Online Social Networks: Unfollowing on Twitter. In *Proceedings of the*

- SIGCHI Conference on Human Factors in Computing Systems* (pp. 1101–1104). New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979105>
- Kivran-Swaine, F., & Naaman, M. (2011). Network Properties and Social Sharing of Emotions in Social Awareness Streams. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (pp. 379–382). New York, NY, USA: ACM. <https://doi.org/10.1145/1958824.1958882>
- Klotz, C., Ross, A., Clark, E., & Martell, C. (2014). Tweet! – And I Can Tell How Many Followers You Have. In S. Boonkrong, H. Unger, & P. Meesad (Eds.), *Recent Advances in Information and Communication Technology* (Vol. 265, pp. 245–253). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-06538-0\\_24](https://doi.org/10.1007/978-3-319-06538-0_24)
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1), 25–36.
- Krathwohl, D. R. (2009). *Methods of educational and social science research: The logic of methods*. Waveland Press.
- Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. Sage.
- Kwak, H., Chun, H., & Moon, S. (2011). Fragile Online Relationship: A First Look at Unfollow Dynamics in Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1091–1100). New York, NY, USA: ACM. <https://doi.org/10.1145/1978942.1979104>
- Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web* (pp. 591–600). ACM.

- Lampe, C. A. C., Ellison, N., & Steinfield, C. (2007). A Familiar Face(Book): Profile Elements As Signals in an Online Social Network. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 435–444). New York, NY, USA: ACM. <https://doi.org/10.1145/1240624.1240695>
- Lange-Faria, W., & Elliot, S. (2012). Understanding the role of social media in destination marketing. *Tourismos. An International Multidisciplinary Journal of Tourism*, 7(1).
- Ledford, C. J. W. (2012). Changing Channels A Theory-Based Guide to Selecting Traditional, New, and Social Media in Strategic Social Marketing. *Social Marketing Quarterly*, 18(3), 175–186. <https://doi.org/10.1177/1524500412460671>
- Lee, J., Ahn, J., Oh, J. S., & Ryu, H. (2015). Mysterious Influential Users in Political Communication on Twitter: Users' Occupation Information and Its Impact on Retweetability. *IConference 2015 Proceedings*. Retrieved from <https://www.ideals.illinois.edu/handle/2142/73456>
- Lengel, R. H., & Daft, R. L. (1989). The selection of communication media as an executive skill. *The Academy of Management Executive (1987-1989)*, 225–232.
- Li, R. (2018). The Secret of Internet Celebrities: A Qualitative Study of Online Opinion Leaders on Weibo. Presented at the Proceedings of the 51st Hawaii International Conference on System Sciences.
- Li, X., Wang, L., & Sung, E. (2005). A study of AdaBoost with SVM based weak learners. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* (Vol. 1, pp. 196–201 vol. 1). <https://doi.org/10.1109/IJCNN.2005.1555829>

- Majchrzak, A., Faraj, S., Kane, G. C., & Azad, B. (2013). The Contradictory Influence of Social Media Affordances on Online Communal Knowledge Sharing. *Journal of Computer-Mediated Communication, 19*(1), 38–55. <https://doi.org/10.1111/jcc4.12030>
- Mansour, O., Askenäs, L., & Ghazawneh, A. (2013). Social Media and Organizing—An Empirical Analysis of the Role of Wiki Affordances in Organizing Practices. Presented at the International Conference on Information Systems International Conference on Information Systems.
- Markus, M. L. (1994). Electronic mail as the medium of managerial choice. *Organization Science, 5*(4), 502–527.
- Markus, M. L., & Silver, M. S. (2008). A foundation for the study of IT effects: A new look at DeSanctis and Poole's concepts of structural features and spirit. *Journal of the Association for Information Systems, 9*(10/11), 609.
- Marshall, P. D. (2006). *The celebrity culture reader*. Routledge.
- Marwick, A. (2013). *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press.
- Marwick, A. (2015a). Instafame: Luxury selfies in the attention economy. *Public Culture, 27*(175), 137–160.
- Marwick, A. (2015b). You May Know Me from YouTube. *A Companion to Celebrity*. Eds. Sean Redmond and P. David Marshall. Chichester: Wiley Blackwell, 333–349.
- Marwick, A., & boyd, danah. (2011). To See and Be Seen: Celebrity Practice on Twitter. *Convergence: The International Journal of Research into New Media Technologies, 17*(2), 139–158. <https://doi.org/10.1177/1354856510394539>



- Marwick, A., & boyd, danah. (2012). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133.  
<https://doi.org/10.1177/1461444810365313>
- Mavroudis, J., & Milne, E. (2016). Researching microcelebrity: Methods, access and labour. *First Monday*, *21*(7).
- McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept (Vol. 2000, pp. 179–186). Presented at the Graphics interface.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*, 415–444.
- McVeigh-Schultz, J., & Baym, N. K. (2015). Thinking of you: Vernacular affordance in the context of the microsocial relationship App, Couple. *Social Media+ Society*, *1*(2), 2056305115604649.
- Metaxas, P. T., Mustafaraj, E., Wong, K., Zeng, L., O’Keefe, M., & Finn, S. (2015). What Do Retweets Indicate? Results from User Survey and Meta-Review of Research. (pp. 658–661). Presented at the ICWSM.
- Montgomery, D. C. (2017). *Design and analysis of experiments*. John wiley & sons.
- Naaman, M. (2012). Social multimedia: highlighting opportunities for search and mining of multimedia data in social media applications. *Multimedia Tools and Applications*, *56*(1), 9–34. <https://doi.org/10.1007/s11042-010-0538-7>
- Naaman, M., Boase, J., & Lai, C.-H. (2010). 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). New York, NY, USA: ACM.  
<https://doi.org/10.1145/1718918.1718953>

- Nagy, P., & Neff, G. (2015). Imagined affordance: Reconstructing a keyword for communication theory. *Social Media+ Society*, 1(2), 2056305115603385.
- Nahon, K., & Hemsley, J. (2013). *Going Viral*. Polity.
- Norman, D. (1988). The Design of Everyday Things (Originally published: The psychology of everyday things). *The Design of Everyday Things (Originally Published: The Psychology of Everyday Things)*.
- Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.
- Nowak, S., & Rüger, S. (2010). How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation. In *Proceedings of the international conference on Multimedia information retrieval* (pp. 557–566). ACM Press.  
<https://doi.org/10.1145/1743384.1743478>
- Opitz, D. W., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *J. Artif. Intell. Res.(JAIR)*, 11, 169–198.
- O’Riordan, S., Feller, J., & Nagle, T. (2012). Exploring the affordances of Social Network sites: an Analysis of Three Networks. (p. 177). Presented at the ECIS.
- Page, R. (2012). The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication*, 6(2), 181–201.
- Papacharissi, Z., & de Fatima Oliveira, M. (2012). Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, 62(2), 266–282.  
<https://doi.org/10.1111/j.1460-2466.2012.01630.x>
- Pegoraro, A. (2010). Look who’s talking—Athletes on Twitter: A case study. *International Journal of Sport Communication*, 3(4), 501–514.

- Pennington, N., Hall, J. A., & Hutchinson, A. (2016). To Tweet or Not to Tweet: Explaining Fan-Celebrity Interaction on Twitter. *Interaction, 48*(1).
- Penrod, J., Preston, D. B., Cain, R. E., & Starks, M. T. (2003). A Discussion of Chain Referral As a Method of Sampling Hard-to-Reach Populations. *Journal of Transcultural Nursing, 14*(2), 100–107. <https://doi.org/10.1177/1043659602250614>
- Petrovic, S., Osborne, M., & Lavrenko, V. (2011). RT to Win! Predicting Message Propagation in Twitter. Presented at the ICWSM.
- Pollach, I. (2008). Media Richness in Online Consumer Interactions: An Exploratory Study of Consumer-Opinion Web Sites. *Information Resources Management Journal, 21*(4), 49–65.
- Pollet, T. V., Roberts, S. G. B., & Dunbar, R. I. M. (2010). Use of Social Network Sites and Instant Messaging Does Not Lead to Increased Offline Social Network Size, or to Emotionally Closer Relationships with Offline Network Members. *Cyberpsychology, Behavior, and Social Networking, 14*(4), 253–258.  
<https://doi.org/10.1089/cyber.2010.0161>
- Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2011). In the Mood for Being Influential on Twitter. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* (pp. 307–314). <https://doi.org/10.1109/PASSAT/SocialCom.2011.27>
- Rahmawan, D. (2013). Selebtwits: Micro-Celebrity Practitioners in Indonesian Twittersphere.
- Ratkiewicz, J., Conover, M., Meiss, M. R., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Detecting and Tracking Political Abuse in Social Media. *ICWSM, 11*, 297–304.

- Rätsch, G., Onoda, T., & Müller, K. R. (1998). An improvement of AdaBoost to avoid overfitting. In *Proc. of the Int. Conf. on Neural Information Processing* (pp. 506–509).
- Raun, T. (2018). Capitalizing intimacy: New subcultural forms of micro-celebrity strategies and affective labour on YouTube. *Convergence*, 24(1), 99–113.
- Reuter, C., Heger, O., & Pipek, V. (2013). Combining real and virtual volunteers through social media. Presented at the Iscram.
- Rice, R. E. (1992). Task Analyzability, Use of New Media, and Effectiveness: A Multi-Site Exploration of Media Richness. *Organization Science*, 3(4), 475–500.
- Robinson, O. C. (2014). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*, 11(1), 25–41.  
<https://doi.org/10.1080/14780887.2013.801543>
- Robson, C., & McCartan, K. (2016). *Real world research*. John Wiley & Sons.
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), 1–39.  
<https://doi.org/10.1007/s10462-009-9124-7>
- Rosenbaum, P. R. (2002). *Observational studies*. Place of publication not identified: Springer.
- Russ, G. S., Daft, R. L., & Lengel, R. H. (1990). Media Selection and Managerial Characteristics in Organizational Communications. *Management Communication Quarterly*, 4(2), 151–175. <https://doi.org/10.1177/0893318990004002002>
- Ryu, D., & Fulk, J. (1991). Group cohesiveness and perceptions of media richness in the workplace (Vol. 23). Presented at the International Communication Association Meeting, Chicago, IL, May.
- Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning*, 5(2), 197–227.  
<https://doi.org/10.1007/BF00116037>

- Schrock, A. R. (2015). Communicative affordances of mobile media: Portability, availability, locatability, and multimodality. *International Journal of Communication*, 9, 18.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering (pp. 815–823). Presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Schwarz, O. (2010). On Friendship, Boobs and the Logic of the Catalogue: Online Self-Portraits as a Means for the Exchange of Capital. *Convergence*, 16(2), 163–183.  
<https://doi.org/10.1177/1354856509357582>
- Seidel, S., Berente, N., Debortoli, S., & Srinivasan, N. (2016). Tweeting Like Taylor Swift? Affordances, Status Production, and Online Platforms. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 1861–1869). IEEE.
- Semaan, B. C., Robertson, S. P., Douglas, S., & Maruyama, M. (2014). Social media supporting political deliberation across multiple public spheres: towards depolarization (pp. 1409–1421). Presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, ACM.
- Semaan, B., Faucett, H., Robertson, S. P., Maruyama, M., & Douglas, S. (2015). Designing political deliberation environments to support interactions in the public sphere (pp. 3167–3176). Presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM.
- Semaan, B., & Hemsley, J. (2015). Maintaining and Creating Social Infrastructures: Towards a Theory of Resilience. Presented at the ISCRAM.
- Senft, T. M. (2008). *Camgirls: Celebrity and community in the age of social networks* (Vol. 4). New York, NY: Peter Lang.

- Sheer, V. C. (2011). Teenagers' Use of MSN Features, Discussion Topics, and Online Friendship Development: The Impact of Media Richness and Communication Control. *Communication Quarterly*, 59(1), 82–103.  
<https://doi.org/10.1080/01463373.2010.525702>
- Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications.
- Simon, S. J., & Peppas, S. C. (2004). An examination of media richness theory in product Web site design: an empirical study. *Info : The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media*, 6(4), 270–281.
- Snickars, P., & Vonderau, P. (Eds.). (2009). *The YouTube reader*. Stockholm: National Library of Sweden.
- Stever, G. S., & Lawson, K. (2013). Twitter as a Way for Celebrities to Communicate with Fans: Implications for the Study of Parasocial Interaction. *North American Journal of Psychology*, 15(2), 339.
- Stieglitz, S., & Dang-Xuan, L. (2012). Political Communication and Influence through Microblogging—An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior. In *2012 45th Hawaii International Conference on System Science (HICSS)* (pp. 3500–3509). <https://doi.org/10.1109/HICSS.2012.476>
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research* (Vol. 15). Newbury Park, CA: Sage.
- Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). 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.

- Suh, K. S. (1999). Impact of communication medium on task performance and satisfaction: an examination of media-richness theory. *Information & Management*, 35(5), 295–312.  
[https://doi.org/10.1016/S0378-7206\(98\)00097-4](https://doi.org/10.1016/S0378-7206(98)00097-4)
- Tanupabrungsun, S. (2017a). IG-Scraper (Version 0.1). Retrieved from  
<https://github.com/stanupab/IG-Scraper>
- Tanupabrungsun, S. (2017b). TW-Scraper (Version 0.1). Retrieved from  
<https://github.com/stanupab/TW-Scraper>
- Tanupabrungsun, S., & Hemsley, J. (in press). Studying Twitter Affordance-Enabled Performance of Celebrity Using a Framework for Measuring Information Richness. *Social Media + Society*.
- Tanupabrungsun, S., Hemsley, J., Semaan, B., & Stromer-Galley, J. (2016). Noisy Candidates and Informative Politicians: Analyzing Changes in Tweet Behavior using Tweet Quality Assessment Framework. iSchools. <https://doi.org/10.9776/16235>
- Thompson, C. (2008). Brave new world of digital intimacy. *The New York Times*, 7.
- Thrall, A. T., Lollo-Fakhreddine, J., Berent, J., Donnelly, L., Herrin, W., Paquette, Z., ... Wyatt, A. (2008). Star power: celebrity advocacy and the evolution of the public sphere. *The International Journal of Press/Politics*, 13(4), 362–385.
- Titlow, J. (2012, July 19). How to Get More Followers on Instagram. Retrieved January 31, 2018, from <https://readwrite.com/2012/07/19/how-to-get-more-instagram-followers/>
- Treem, J. W., & Leonardi, P. M. (2013). Social media use in organizations: Exploring the affordances of visibility, editability, persistence, and association. *Annals of the International Communication Association*, 36(1), 143–189.

- Trevino, L. K., Lengel, R. H., & Daft, R. L. (1987). Media Symbolism, Media Richness, and Media Choice in Organizations A Symbolic Interactionist Perspective. *Communication Research, 14*(5), 553–574. <https://doi.org/10.1177/009365087014005006>
- Tsikerdekis, M., & Zeadally, S. (2014). Online Deception in Social Media. *Commun. ACM, 57*(9), 72–80. <https://doi.org/10.1145/2629612>
- Turner, G. (2004). *Understanding Celebrity* (First Edition). SAGE Publications.
- Turner, G. (2010). *Ordinary people and the media: The demotic turn*. Sage Publications.
- Turner, G. (2013). *Understanding Celebrity* (Second Edition). SAGE Publications.
- Twitter, Inc. (2017). About Twitter, Inc. | About. Retrieved March 2, 2016, from <https://about.twitter.com/company>
- Usher, B. (2015). Twitter and the celebrity interview. *Celebrity Studies, 6*(3), 306–321.
- Uysal, I., & Croft, W. B. (2011). User oriented tweet ranking: a filtering approach to microblogs. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 2261–2264). ACM.
- Vadivu, V. M., & Neelamalar, M. (2015). Digital brand management—A study on the factors affecting customers' engagement in Facebook pages (pp. 71–75). Presented at the Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015 International Conference on, IEEE.
- Walther, J. B. (1992). Interpersonal Effects in Computer-Mediated Interaction A Relational Perspective. *Communication Research, 19*(1), 52–90. <https://doi.org/10.1177/009365092019001003>



- Wang, Q., Woo, H. L., & Quek, C. L. (2012). Exploring the affordances of Facebook for teaching and learning. *International Review of Contemporary Learning Research, 1*(1), 23–32.
- Wang, Y., & Kraut, R. (2012). Twitter and the Development of an Audience: Those Who Stay on Topic Thrive! In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1515–1518). New York, NY, USA: ACM.  
<https://doi.org/10.1145/2207676.2208614>
- Wang, Y., Li, Y., Semaan, B., & Tang, J. (2016). Space Collapse: Reinforcing, Reconfiguring and Enhancing Chinese Social Practices through WeChat. (pp. 427–435). Presented at the ICWSM.
- Ward, J. (2016). A Content Analysis of Celebrity Instagram Posts and Parasocial Interaction. *Elon Journal of Undergraduate Research in Communications, 7*(1). Retrieved from <http://www.inquiriesjournal.com/articles/1478/a-content-analysis-of-celebrity-instagram-posts-and-parasocial-interaction>
- Wellman, B. (2001). Physical Place and Cyberplace: The Rise of Personalized Networking. *International Journal of Urban and Regional Research, 25*(2), 227–252.  
<https://doi.org/10.1111/1468-2427.00309>
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2003). The social affordances of the Internet for networked individualism. *Journal of Computer-Mediated Communication, 8*(3), 0–0.
- Weng, J., Lim, E.-P., Jiang, J., & He, Q. (2010). TwitterRank: Finding Topic-sensitive Influential Twitterers. In *Proceedings of the Third ACM International Conference on*

- Web Search and Data Mining* (pp. 261–270). New York, NY, USA: ACM.  
<https://doi.org/10.1145/1718487.1718520>
- Wikipedia contributors. (2016, May 16). Hashtag. In *Wikipedia, the free encyclopedia*. Retrieved from <https://en.wikipedia.org/w/index.php?title=Hashtag&oldid=720504608>
- Wikipedia contributors. (2017, March 12). List of most followed users on Instagram. In *Wikipedia, the free encyclopedia*. Retrieved from [https://en.wikipedia.org/w/index.php?title=List\\_of\\_most\\_followed\\_users\\_on\\_Instagram&oldid=770014567](https://en.wikipedia.org/w/index.php?title=List_of_most_followed_users_on_Instagram&oldid=770014567)
- Xu, B., Huang, Y., Kwak, H., & Contractor, N. (2013). Structures of broken ties: exploring unfollow behavior on twitter (p. 871). ACM Press.  
<https://doi.org/10.1145/2441776.2441875>
- Yang, Z., Guo, J., Cai, K., Tang, J., Li, J., Zhang, L., & Su, Z. (2010). Understanding retweeting behaviors in social networks. In *ACM Press* (p. 1633). ACM Press.  
<https://doi.org/10.1145/1871437.1871691>
- Yanow, D., & Schwartz-Shea, P. (2015). *Interpretation and method: Empirical research methods and the interpretive turn*. Routledge.
- Zaman, T. R., Herbrich, R., Van Gael, J., & Stern, D. (2010). Predicting information spreading in twitter (Vol. 104, pp. 17599–17601). Presented at the Workshop on computational social science and the wisdom of crowds, nips, Citeseer.
- Zappavigna, M. (2017). Twitter. In *Pragmatics of Social Media* (Vol. 11, pp. 201–224).
- Zheng, Y., & Yu, A. (2016). Affordances of social media in collective action: the case of Free Lunch for Children in China. *Information Systems Journal*, 26(3), 289–313.

## VITA

### Sikana Tanupabrungsun

PhD Candidate  
 School of Information Studies  
 Syracuse University  
<http://my.ischool.syr.edu/People/stanupab>  
[stanupab@syr.edu](mailto:stanupab@syr.edu)

### EDUCATION

2018 ( <i>Expected</i> )	Ph.D.	Information School, Syracuse University
2014	M.Eng.	Computer Engineering King Mongkut's University of Technology Thonburi, Thailand
2009	B.Eng.	Computer Engineering King Mongkut's University of Technology Thonburi, Thailand

### EMPLOYMENT

8/14 to present *Research Assistant*, School of Information Studies, Syracuse University,  
 Syracuse, NY  
 Research: Research focuses on online behaviors through analysis of social media data using mixed methods.

6/11 to 7/14: Software Developer, Software Computing and Innovation Center, Bangkok, Thailand

8/12 to 10/12: Product Specialist Intern, Microsoft, Bangkok, Thailand

8/11 to 10/11: Research Intern, Shibaura Institute of Technology, Tokyo, Japan

### SCHOLARSHIP

#### Journals and Annuals (3)

Tanupabrungsun, S., Hemsley, J. (in press). Studying Twitter Affordance-Enabled Performance of Celebrity Using a Framework for Measuring Information Richness. *Social Media + Society*.

Tanupabrungsun, S., Hemsley, J., Semaan, B. (2018). Information affordances: Studying the information processing activities of the core Occupy actors on Twitter. *First Monday*, 23(2).

Hemsley, J., Stromer-Galley, J., Tanupabrungsun, S., Semaan, B. (2018) Tweeting to the Target: Candidates Use of Strategic Messages and @Mentions on Twitter. *Journal of Information Technology & Politics*, 15(1).

#### Refereed Conference Proceedings (12)

Hemsley, J., S. Tanupabrungsun. 2018. Dribbble: Exploring the Concept of Viral Events on an Art World Social Network Site. Springer Lecture Notes, proceedings of the iConference, Sheffield, UK. (30% acceptance)

Boichak, O., S. Jackson, J. Hemsley, S. Tanupabrungsun. 2018. Automated Diffusion? Bots and their Influence During the 2016 U.S. Presidential Election. Springer Lecture Notes, Proceedings of the iConference, Sheffield, UK. (30% acceptance)

Hemsley, J., Tanupabrungsun, S., & Semaan, B. (2017). Call to Retweet: Negotiated diffusion of strategic political messages. In Proceedings of the 8th International Conference on Social Media & Society. Toronto: ACM.

- Rossini, P., Hemsley, J., Tanupabrungsun, S., Zhang, F., Robinson, J., & Stromer-Galley, J. (2017). Social Media Strategies and Public Opinion Polls in the Early Stages of the 2016 U.S. Presidential Campaigns. In Proceedings of the 8th International Conference on Social Media & Society. Toronto: ACM.
- Zhang, F., Tanupabrungsun, S., Hemsley, J., Robinson, J., Semaan, B., Bryant, L., Stromer-Galley, J., Boichak, O., Hegde, Y. (2017). Strategic Temporality on Social Media During the General Election of the 2016 U.S. Presidential Campaign. In Proceedings of the 8th International Conference on Social Media & Society. Toronto: ACM.
- Zhang F., Stromer-Galley J., Tanupabrungsun S., Hegde Y., McCracken N., Hemsley J. (2017) Understanding Discourse Acts: Political Campaign Messages Classification on Facebook and Twitter. In: Lee D., Lin YR., Osgood N., Thomson R. (eds) Social, Cultural, and Behavioral Modeling. SBP-BRiMS 2017. Lecture Notes in Computer Science, vol 10354. Springer, Cham
- Stromer-Galley, J.; Zhang, F.; Tanupabrungsun, S.; Hemsley, J.; Semaan, B., 2016. Tweeting the attack: Predicting gubernatorial candidate attack messaging and its spread. 66th annual conference of the International Communication Association (ICA), Fukuoka, Japan (46% acceptance)
- Tanupabrungsun, S., J. Hemsley, B. Semaan, J. Stromer-Galley, 2016. Noisy Candidates and Informative Politicians: Analyzing Changes in Tweet Behavior using Tweet Quality Assessment Framework. Proceedings of the iConference, Philadelphia, PA. (30% acceptance) *\*Finalist for Lee Dirks Award for Best Paper*
- Tanupabrungsun S. and T. Achalakul, Feature Reduction for Anomaly Detection in Manufacturing with MapReduce GA/kNN, Proceedings of the 19th IEEE International Conference on Parallel and Distributed Systems (ICPADS's Workshop on SC-BDA), 15-18 December 2013, Soul, Korea. *\*Best Student Paper Award*
- Pattananagkur T., Tanupabrungsun S., Areekijseere K., Pumma S., and Achalakul T., The Design of SkyPACS: a High Performance, Mobile Medical Imaging Solution, Symposium on GPU Computing and Applications, 9 October 2013, Singapore.
- Tivatansakul S., Tanupabrungsun S., Areekijseere K., Achalakul T. and Ohkura M., Intelligent Space Design for the elderly. in proceedings of the 6th International Symposium, South East Asian Technical University Consortium (SEATUC2012), 6-7 March 2012.
- Tivatansakul S., Tanupabrungsun S., Areekijseere K., Achalakul T., Hirasawa K., Sawada S., and Ohkura M., The intelligent space for the elderly—Implementation of fall detection algorithm. Proceedings of the SICE Annual Conference, 2012, 1944-1949. IEEE, 2012.

#### Workshops (1)

Hemphill, Libby; Rogers, Richard; Hemsley, Jeff; Tanupabrungsun, Sikana. 2017. Social Media Data Bootcamp and Research Hackathon. Internet Research 18: The 18th Annual Meeting of the Association of Internet Researchers. Tartu, Estonia, October, 2017.

#### Conference Extended Abstracts and Presentations (4)

- Hemsley, Jeff; Sonne, Jennifer; Tanupabrungsun, Sikana; Yu, Yihan; Deekshitula, Suchitra. 2017. A Study Of Diffusion In The Dribbble Art World. Internet Research 18: The 18th Annual Meeting of the Association of Internet Researchers. Tartu, Estonia, October, 2017.
- Boichak, Olga; Hemsley, Jeff; Tanupabrungsun, Sikana; Espinosa, Daniela Fernandez; Jackson, Sam; Rossini, Patricia. 2017. Automated Diffusion? Bots And Their Influence During The 2016 U.S. Presidential Election. Internet Research 18: The 18th Annual Meeting of the Association of Internet Researchers. Tartu, Estonia, October, 2017.
- Jennifer Stromer-Galley, Sikana Tanupabrungsun, Feifei Zhang, Jeff Hemsley, Nancy McCracken, Lauren Bryant, Brian Dobreski, Bryan Semaan. 2016. Strategic Changes in Gubernatorial Campaign Messaging Over Time: A Computational Analysis. 2016 International Conference on Computational Social Science. Evanston, IL

Hemsley, J., S. Tanupabrungsun, B. Semaan, J. Stromer-Galley, 2015. Attack and Position: U.S. Gubernatorial Candidate Direct Campaign Dialogue on Twitter. *Internet Research 16: The 16th Annual Meeting of the Association of Internet Researchers*. Phoenix, AZ, October, 2015.

#### Refereed Posters (5)

- Hemsley, J., Suchitra, D., Tanupabrungsun, S., Sonne, J. Yu, Y. 2017. A Study of Diffusion in the Dribbble Art World. 8th International Conference on Social Media & Society. Toronto
- Boichak, O., Tanupabrungsun, S., Espinosa, D.F., Hemsley, J. 2017. Automated Diffusion? Bots and Their Influence During the 2016 U.S. Presidential Election. 8th International Conference on Social Media & Society. Toronto
- Tanupabrungsun, S., Hemsley, J. 2017. Microcelebrity Practices: Towards Cross-Platform Studies Through a Richness Framework. 8th International Conference on Social Media & Society. Toronto
- Tanupabrungsun, S., J. Hemsley, B. Semaan, J. Stromer-Galley, 2016. GOPDebate: What the Public is "Talking" About on Instagram. Proceedings of the iConference, Philadelphia, PA.
- Tanupabrungsun, S., J. Hemsley, B. Semaan, J. Stromer-Galley, 2015. Gubernatorial @mentioning Message Strategy on Twitter. Social Media & Society Conference, Toronto, Canada.

#### Software (1)

Hemsley, J., Ceskavich, B., Tanupabrungsun, S. (2014). STACK (Version 1.0). Syracuse University, School of Information Studies. Retrieved from <https://github.com/jhemsley/Syr-SM-Collection-Toolkit>. DOI: 10.5281/zenodo.12388.

## **GRANTS AND AWARDS**

#### Grants (1)

Sikana Tanupabrungsun. 2017. Microcelebrity Practices: Towards Cross-Platform Studies Through a Scoring Framework. Center for Computational and Data Sciences Seed Funding Award, \$1,125.00

#### Awards and Honors

- 2017 Doctoral Colloquium Award, iConference, Wuhan, China.
- 2016 Finalist for Lee Dirks Award for Best Paper, iConference, Philadelphia, PA.
- 2016 CRA-W Grad Cohort Award, San Diego.
- 2013 Best Student Paper award, ICPADS workshop on Scalable Computing for Big Data Analytics, Seoul, Korea.
- 2013 Winner of the Thailand Imagine Cup 2013 (Innovation: SkyPACS).
- 2013 Third place winner of the Imagine Cup Worldwide Final 2013, Russia (Innovation: SkyPACS).
- 2013 Winner of the National Software Contest 2013, Thailand (Science and Technology).
- 2013 Winner of the Thailand ICT Awards 2013.
- 2012 Winner of the KMUTT medal for the academic achievement (undergraduate).
- 2012 Winner of the Thailand Imagine Cup 2012 (Software Design: The Smart House).
- 2012 Winner of the Student Design Challenge (iCREATe) 2012, Singapore.
- 2012 TRF-MAG Center Scholarship for graduate study, Bangkok, Thailand.

**SERVICE**Institutional Academic Service

Syracuse University

2014/2015

PhD Student Representative, PhD Committee

Reviewing

Conferences: iConference, Association of Internet Researchers (AoIR), ACM CHI, Social Media and Society International Conference (SMS), International Conference on Information Systems (ICIS)