

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

A Library of Fragments: Digital Quotations, New Literacies, and Attention on Social Media

### Permalink

<https://escholarship.org/uc/item/7nj2j0jb>

### Author

Booten, Kyle Paul

### Publication Date

2017

Peer reviewed|Thesis/dissertation

A Library of Fragments:  
Digital Quotations, New Literacies, and Attention on Social Media

by

Kyle Paul Booten

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Education

and the Designated Emphasis

in

New Media

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Sarah Warshauer Freedman

Professor Marti A. Hearst

Professor Glynda Hull

Spring 2017

A Library of Fragments: Digital Quotations, New Literacies, and Attention on Social Media

© Kyle Booten, 2017

## Abstract

### A Library of Fragments: Digital Quotations, New Literacies, and Attention on Social Media

by

Kyle Paul Booten

Doctor of Philosophy in Education

Designated Emphasis in New Media

University of California, Berkeley

Professor Sarah Warshauer Freedman, Chair

From tweets to GIFs to memes, social media is awash in bite-sized texts that are perfect for quick, instantaneous consumption and viral sharing. Mixed up in the raucous frenzy of social media, however, are excerpts that originate from print media: quotations from novelists, poets, philosophers, and other authors. The book, far from extinct, has nevertheless become "fragmented," circulating in new ways on digital networks.

This study examines the phenomenon of quotations—especially those from books—on the social network Tumblr. Its fundamental question is whether the fragmentation of the book represents a threat to traditional forms of attentive, immersive engagement with long-form texts.

The study reports on three research phases, each with distinct methods and data. The first phase, relying on qualitative and quantitative discourse analysis, examined the role of books and traditionally print-based discourses within the context of Tumblr. The second employed ethnographic data techniques, questionnaires and interviews, to uncover Tumblr users' purposes for sharing quotes from books as well as the ways that quoting connects to these users' broader experiences of literacy. The third phase used computational linguistics to distinguish between different types of quotations and to illuminate the features that contribute to "quotability."

This study finds that quotation can be—though is not always—intertwined with traditional forms of reading and literate attention. Chapter 3 demonstrates that the quoting of books in particular is a fundamental part of the larger practice of quoting on Tumblr; it further illustrates some of the wide variety of quotations that circulate on this network, from misquotations attributed to famous authors to those quotes that are carefully annotated with bibliographic information, scholarly vocabulary, and other signs of deep familiarity with source texts. Chapter 4 reveals that, for those who quote from books on Tumblr, quoting is often not a form of distraction; rather it can deepen their experience of reading and lead them to discover new books. Yet quoting also serves more fundamental purposes of self-expression and self-care, functions that may be related to the fact that quoting is a starkly gendered (largely female) practice. Chapter 5 suggests that,

while many quotations are indeed designed to be easily-decontextualized adages, others are specifically tailored to fulfill quoter's desires to express their intimate thoughts and feelings. Some quotes seem to be tailored to be appreciated by those with deep familiarity with the source text, though such quotations are less popular than easily-decontextualized ones.

The fragmentation of the book does not fundamentally threaten traditional forms of reading, though the very purposes of reading are increasingly bound up with self-expression and self-care made possible through social media. The conclusion chapter considers the ramifications of these findings for media studies, the history of the book, and educational practice.

## Acknowledgements

First, I express my gratitude to Sarah Freedman, who has been an unfailingly wise mentor, and who has modeled a way of approaching the work of research with equal parts passion and purpose. I hope to become a scholar who moves as fluently as she does between the tantalizing abstractions of theory and complex and energizing textures of everyday life. To Marti Hearst, whose patience and encouragement have allowed me stretch far beyond my comfort zones, and who has challenged me to become my own kind of programmer. And to Glynda Hull, whose creativity continues to inspire me; how lucky I was to find my place among the collection of people and projects that she has assembled. Together, these figures helped me navigate across and between disciplines, finding a path that was—if not the straightforward one—the right one.

To other key mentors in the Graduate School of Education, the Berkeley Center for New Media, and beyond : Laura Sterponi, Kris Gutiérrez, Patricia Baquedano-López, Greg Niemeyer, Abigail De Kosnik, Robert Glushko, Shirley Brice Heath, and Paige Ware. The pages that follow reverberate with the echoes of their thoughtful critiques and their generous conversation. And to Barbara White: I remember with thanks her insight and her kindness. To Lara Wolfe, Billie Jo Conlee, Karen Sullivan, Rosa Garcia, and Ilka Williams: a university like Berkeley can seem vast and labyrinthine, but, because of them, this was not my experience. I am grateful to have had them as guides.

To friends from the Bay: Amy Koehler Catterson, Daniel Catterson, Jen DiZio, Emily Hellmich, Keith Martin, John Scott, Jessica Xiomara Garcia, Camilo Landau, Haley Pollack, Michael David Lukas, and Herman Leung. This project would have long ago sunk beneath its own ponderousness were it not for the levity provided by their wit. Special thanks to Jenni Higgs, my comrade. And thanks to old friends: Taylor Woodroof, Michael Gorwitz, and Adam Hayden. Their friendship has often reminded me of who I once was and who I might yet be.

To Patricia Lutz and Pamela Costa, who in different ways helped me get through some of the most challenging moments of this process. And to Ann Booten, for her *joie de vivre*.

To Don and Paula Booten: I have been blessed with parents who have supported me unconditionally in countless ways across many years of school and through crises large and small. Their encouragement and their humor made this dissertation possible, and so I dedicate it to them.

And thanks to Emma, my love: her presence in my life sustained me through long seasons of research, coding, and writing. I am constantly learning not just from her conversation but from the way she approaches life with fearlessness and joy. And to her parents, Salvador Amador and Deborah Reger, for always making me feel at home in their adventure.

Thanks to the Tumblr users who helped me to understand their world and whose perspectives constantly surprised me. Finally, I express my thanks to the National Academic of Education/Spencer Dissertation Fellowship Program for their generous support. And I thank the Berkeley Center for New Media for supporting me with the Peter Lyman Fellowship and a Summer Research Award.

## Contents

<b>Chapter 1: Introduction</b>	1
Review of the Literature	2
A Cognitive Crisis?	2
The Fate of Books and Reading	3
Social Media, Literacy, and Literature	5
Other quantitative studies of quotability	7
"Geeking Out" Online	8
Social Media and Audience	9
Gaps in the Literature	10
Theoretical (and Historical) Framing	10
Sharable Media	10
Appropriation	12
Different Views on Practice	13
Commonplace Books	14
Research Questions	15
<b>Chapter 2: Methodological Introduction: Computational Textual Ethnography</b>	17
Why Tumblr?	17
From "Virtual Ethnography" to "Computational Textual Ethnography"	20
On APIs, Hashtags, and Computation	23
On Analytic Units	24
On Words	25
What "Books"?	25
Phases of Research	25
<b>Chapter 3: The Ghosts of Old Media on Tumblr</b>	26
Data and Methods	26
Data	27
Hashtag Corpora.	27
Sources Corpora.	27
Simple BQ Corpus.	28
Philosopher Quote Corpus/Hashtag Corpus.	28
BQ Photo Corpus.	28
Possible misquotes.	29
Usernames Corpus/Blog Description Corpus.	29
Audience Usernames Corpus.	29
A focal quote.	30
Analysis	30
Quantifying tags/visualizing tag relationships.	30
Analyzing Sources Corpora.	31
Analyzing Simple BQ Corpus.	31
Zooming in on philosophic tags.	31
Analysis of photo-type posts.	31

Checking for misquotes.	32
Analyzing usernames/blog description corpus.	32
Analyzing Audience Usernames Corpus.	32
Analyzing focal quote.	32
Findings	33
Old Media, Including Books, are an Important Part of Quotation Practice in General	33
Book Quotes Are Different Than Quotes in General In Certain Ways (And Not In Others)	38
Hashtags Suggest "Deep" ...and Less "Deep" ...Forms of Literate Attention A note on the ambiguity of tags.	40
Quoters Use Academic Vocabulary	41
Multimodal Quotations	46
Quotation celebrates the book as a medium.	46
Quotations are also transmedia artifacts.	52
Other Signs of Practice	53
Misquotation.	53
Misquotations do occur on Tumblr.	53
Vernacular bibliography: misquotation's antidote.	55
Quotations Within Blogs.	56
"Bookish" Blogs.	56
Bookish blogs vs. non-bookish audiences.	56
Discussion	58
<b>Chapter 4: Quoters in their Own Words</b>	60
Data and Methods	60
Data	60
Questionnaire: Reaching Out to Tumblr Users.	60
Participants.	60
Follow-up interviews.	61
Blog Description Corpus.	62
Analysis	62
Analyzing questionnaire responses.	62
Analyzing follow-up interviews.	63
Analyzing Blog Description Corpus.	63
Exploring "nonhuman" Tumblr users.	63
Findings	63
Who Quotes?	63
Gender and age.	64
Theorizing Quotability	65
Powerful language.	65
Semantic aspects, especially "resonance."	66
Why Quote?	67
Quoting is a form of self-expression.	67
Self-expression can be a form of self-care.	68
"Inspiring" quotations and other-care.	69



Quotation is a gendered literacy practice.	70
Quoting and books.	71
Quotations and bibliophilia.	71
Quotations as advertisements.	71
A note on corporate activity.	72
John Green's "shtick."	73
Modalities of Attention	75
Connection to composition.	79
Quotations and memory.	80
Discussion	82
<b>Chapter 5: Analyzing the Linguistic Character of Book Quotes</b>	<b>83</b>
Methodological Overture	83
Distant and Close Linguistic Analysis	83
Linguistics and "Computational Textual Ethnography."	84
Literary language.	84
Data, Features, and Methods	85
Data	85
Tumblr Data	86
Book Quotes (BQ)/General Quotes (GQ) Corpora.	86
BQ 1-Sent/2-Sent Corpora.	86
Pop/Unpop Corpus.	86
BQ Pop/Unpop Blogs Corpora.	86
Comparison Corpora.	87
Brown Corpus.	87
Brown Lit Corpus.	87
Brown Lit 1-Sent Corpus.	87
Twitter Quotes Corpus.	87
Features	88
Tokens.	88
Subject Nouns.	88
High-level syntax patterns.	88
Length.	89
Person names.	89
Imperatives.	89
Repetition.	89
Analyses	89
Exploratory comparisons.	90
Analyzing examples.	90
Hypothesis testing.	91
Further Analyses.	91
Findings	92
Length	92
Key Words and Key Subjects	92
Big words for big questions.	94
Everything or nothing.	95

Copular "is."	95
Emotions.	96
Pronouns.	97
Gendered vs.	
non-gendered personal pronouns.	98
Key Syntax	99
Other discursive patterns.	101
"Just."	102
"There."	103
Testing specific hypotheses	103
BQ 1-Sent Corpus vs. Brown Lit 1-Sent Corpus.	103
Pop Quotes Corpus vs. Unpop Quotes Corpus.	103
A caveat: whose "good" quote?	103
Blog-Level Trends: Popular vs. Unpopular Blogs	106
The Role of the Author	106
Are Quotations Network Specific?	107
Discussion	108
<b>Chapter 6: In Conclusion</b>	110
Summary of the Study	110
Limitations of the Study	111
Ethnography and Computation	111
Implications	112
Beyond Hyper and Deep	112
"Social" Reading	113
Beyond Bibliophilia	114
Lingering and Emergent Questions	115
Interdisciplinary Connections	116
Quoting on Tumblr: Is "New media" new?	116
Scholarly vs. Vernacular Practice.	117
Self-Care.	118
Educational Implications.	118
Possible connections between quoting	
and school-based literacy.	119
Critical versus "uncritical" reading.	119
Towards a New Rhetoric.	120
References	121
Appendices	130

## CHAPTER 1 INTRODUCTION

The student's stylus digs into the wax tablet. The scroll unfurls in the unsteady hand of the novice orator. Puzzled, a young scholar carefully flips back and forth the pages of a codex. The scribe, copying an ancient text, learns to "fix" certain infelicities as he goes along. A young girl, hiding in her room from her governess, escapes into the world of a printed novel.

The history of education is a history of changing technologies. But it is also the story of the tensions between the official, school-sanctioned uses of these technologies and their popular, demotic uses. Today, the divide is greater than ever. One key way in which "new media" are new is the extent to which they have proliferated *outside* of traditional educational institutions, which for so long could claim primacy (if not outright hegemony) in the worlds of reading and writing. Beyond the boundaries of schools, on their laptops and their phones, young people produce and consume texts in ways that look very different than traditional classroom literacy practice (Ito et al., 2009). Devoted to possibly outdated definitions of what counts as reading and writing, schools run the risk of alienating students (Buckingham, 2007; Hayles, 2007).

As media theorist Katherine Hayles (2007) argues, perhaps the most important way that young peoples' outside digital literacies depart from their in-school literacies is in the modes of attention that they require. Schools have traditionally been places for the cultivation of "deep attention," sustained focus to one or a small number of texts. Compare this to the student who engages in "hyper attention," flipping anxiously and erratically between ten or twenty browser tabs—one showing an online problem set, others buzzing with various social networking feeds, YouTube videos, blogs, still others containing half-typed emails or required "reading responses" for another class. Such a change in styles of attention has consequences for the ways young people learn. As Carr (2011) reports, mounting proof from the field of cognitive science suggests that hypertext and web technologies more broadly may increase the cognitive demands of processing information and contribute to distraction and shallow understanding of material. What are schools to do in the face of these changing regimes of cognition and attention? Should they serve as cultural bulwarks against a general slide into "digital distraction?" Or should they, as Hayles (2007, p. 198) has suggested, explore the "frustrating, zesty, and intriguing ways" that "hyper" and "deep" attention may interact?

Researchers, policymakers, and educators have a vital stake in understanding the myriad ways that reading is changing in the context of digital media. As a way of accounting for possible changes with special focus given to the important matter of attention, this study explores a significant aspect of digital culture that, until this point, has generally not been considered by literacy researchers. While popular and scholarly discourse frames social networks as places where people (over)share their own musings and opinions (Senft, 2013), some social media users spend a great deal of time sharing primarily the words of others—in other words, they *quote*. Pinterest, Tumblr, Twitter, Reddit, and of course Facebook are all home to robust cultures of quotation. Members of these sites painstakingly retype quotations from books, speeches, and songs. Noticing the importance of quotation on social networks, the networks themselves have actively encouraged it; Tumblr, for instance, from the start offered users the option of using a "quote" data-type, while Pinterest now offers new users the option of following categories like "inspiring quotes" upon signing up.

In this dissertation, I conceptualize online quotation culture as a "new literacy."

According to Lankshear and Knobel's (2007) definition, for a literacy to be new, it must be novel both in terms of its technical aspects and what they call its "ethical" aspects: a new literacy must be practiced according to new norms and values, and it should give rise to different ways of thinking and being in the world. Following Hayles (2007), I focus on the concept of *attention* as itself a key term for the understanding of this new literacy. This phenomenon is a unique example of the transformations that are occurring in the ways that digitally-networked people are reading and writing. If the novel is the quintessential genre of "deep attention," short texts like the quotation may be a key aspect in the future of reading--a literate attention devoted not to any one text but scattered over or threaded among a multitude of textual fragments whose meaning and beauty can be grasped instantaneously. In fact, since many quotations originate in books, studying online quotation culture also can help chart the fate of the book in a digital age. While there is much that is both technically and ethically novel about online quotation culture, it is also an object lesson in the continuing negotiations between old and new ways of reading and writing and possibilities for hyper and deep attention online.

## **Review of the Literature**

### **A Cognitive Crisis?**

Evidence is mounting that we are in fact in the midst of a generational shift in cognitive styles brought about by the mass proliferation of digital technologies, and nowhere can these changes be seen more readily than in the lives of young people. Studies by the Kaiser Foundation have charted an important shift in the ways youth employ technology toward the use of multiple media simultaneously—for instance, sending emails while watching TV, or carrying on a conversation via text-message while also reading a book (Rideout, Foehr, & Roberts, 2010). An experimental study of people using multiple media (computers and television) simultaneously found that, compared with adults, young people switch between media sources more frequently (Brasel & Gips, 2011). In a recent study, young people engaging in multitasking on computers maintained focused on a single task for less than six minutes (L. D. Rosen, Carrier, & Cheever, 2013). The shift in styles of attention impacts the ways that young people learn, though researchers are still sorting out exactly how. A growing body of research suggests that youth use technology in ways that could be considered "hyper attentive" or distracted, including using social media during homework (Junco & Cotten, 2012) and class (Kraushaar & Novak, 2010; Wood et al., 2012), hurting academic performance (see also Jacobsen & Forste, 2011; Kirschner & Karpinski, 2010). While studies have found that multitasking does not negatively impact reading comprehension (Fox, Rosen, & Crawford, 2009; Lin, Robertson, & Lee, 2009), Levine et al. (2007) found subjective measures of "distractibility" to correlate positively with the frequency of using digital technology (instant messaging) and negatively with the frequency of reading books. These and other findings, they suggest, are consistent with the position that "IMing helps to create a cognitive style based on quick, superficial multitasking rather than in-depth focus on one task such as reading" (pp. 565).

These complex and ongoing changes represent a clear challenge to those institutions that have historically been charged with developing the literate capacities of young people: schools. Across higher education, there is a growing sense that teachers must shift expectations to accommodate the hyper-attention of students by, for instance, thinking twice before assigning lengthy tomes, as well as to advocate for the power of deeply attentive experiences with novels and other texts (Hayles, 2007). It may also be the true, as Hayles suggests, that the shift from deep to hyper literacy is bringing about its own affordances. In fact, certain forms of online

social networking, which can be scorned as mere distractions, may actually help students in school; Junco (2012) found that in particular sharing links to blogs, news articles, and other texts was predictive of improved academic performance in a group of college students, perhaps because sharing links itself resembles the sorts of information perusing and sharing that are important in academic work. And certainly there may be other cognitive effects wrought by digital media use. Sparrow, Liu, and Wegner (2011) suggested that the ubiquity of the internet does change the way we think, placing less importance on remembering information and greater emphasis on where information can be found (i.e., how to Google it). In this sense, memory is not becoming worse or better but different, tailored to interactions with digital media.

This discussion of the consequences of the changing dynamics of literacy is reminiscent of Vygotsky's (1980, 2012) theorization of the ways that the symbolic and material tools historically co-evolve with the mind. Tools—including the myriad technologies of literacy, from the wax tablet to the search engine—can give rise to new "higher mental functions." New ways of thinking can in turn give rise to the creation of new tools. A comparison could also be made to classic anthropological debates about the cognitive consequences of literacy (see Freedman, Hull, Higgs, & Booten, 2016). While some scholars suggested that the invention of literacy itself led to vast cognitive differences between literate and non-literate peoples (Havelock, 1986; Ong, 1982/2013), research revealed this to be a gross exaggeration. At best, the transition from non-literate to literate culture brings about certain limited cognitive changes that are specific to particular literacy practices; for instance, Scribner and Cole (1981) found that long-term practice of memorizing the Quran led to modest improvement of certain kinds of memory. With this in mind, views that cast the cognitive repercussions of social media in extreme terms (whether positive or negative) likely overstate the case. Moreover, it is important not to measure new technologies against the yardsticks inherited from previous literacy technologies. Insofar as schools (and other institutions) still value deep attention and judge students according to their capacity for it, it is no wonder that digital media that demand hyper-attention generally do not seem to contribute to school success.

While cognitive research on changing styles of attention is no doubt crucial and invaluable, a Vygotskian perspective on the socio-historical development of mental capacities suggests that it is equally important to consider in a culturally-attuned and historically-sensitive way the developing tools and systems that solicit our attention in different ways—especially *the book*.

### **The Fate of Books and Reading**

Hayles' worries about the distinctions between "deep" and "hyper" take part in a larger discussion of what is happening to old media, especially the book, during the age of the book. These discussions largely occur within the humanities, including media studies and literary studies, and so discussions about the "fate of the book" often blur into those about the "fate of literature."

Even as she worries about the fate of traditional forms of literate attention, Hayles is measured in her evaluation of the ways that "hyper" attention has its affordances. This is not true of Sven Birkerts (2006); in his *Gutenberg Elegies* (the title is worth mentioning, as it is often referenced in later works) Birkerts painted the rise of digital technology, with its tendency to distract, as an unmitigated disaster for literacy and literature. For Birkerts, the fact that the contemporary reader "tends to move across surfaces, skimming, hastening" precludes a profound engagement with literature in which words "resonate inwardly" (pp. 72). In this context, Birkerts

suggests, the book as an old medium actually takes on a sort of politically significant reactionary potential. It is a "counter-technology" — "counter" in the sense that it is so much at odds with increasingly frenetic and "interactive" new media. Birkerts also noted that the rise of digital technology seemed to signal a loss of capacity for memorization (a skill he attributed to societies that suffer from a dearth rather than an overabundance of texts).

Opposed to Birkerts and the techno-pessimism he represents are those theorists who have imagined that the book itself is not being erased by digital media but rather transformed by it, in some ways for the better. In his enduring and well-known vision of a post-book utopia, Nelson (1993) imagined that soon the book — as a stable, consistent, and often single-authored artifact — would be replaced by unstable, constantly-evolving digital texts compiled on-the-fly from the contributions of multiple authors. Likewise, Landow (1991) saw in "hypertext" the potential for new (again, mostly authorless) forms of writing as well as reading. He valorized emergent genres of narrative fiction that eschewed any sort of linear narrative, instead leaving the reader to interactively wander through a web of interlinking narrative fragments. Bolter (1996) suggested that, contrary to Ong and McLuhan, who saw the rise of digital technology as a return to orality, digital texts were increasingly visual. Lingering below these imaginations from the early to mid 1990s of what the book would shortly become are admissions that the kind of deep, linearity traditionally expected by books would have to be replaced by different modes of attention.

Studies of the future of the book, literature, and reading have often eschewed both dour pessimism and techno-utopianism in favor of a more measured position. According to Thornberg and Jenkins (2003), "[o]ld media rarely die," complex synergies...always prevail among media systems, particularly during periods shaped by the birth of a new medium of expression" (pp. 3). This perspective has been elaborated with respect to the book in particular by Striphos (2009). His study considered the ways that the book (including the *printed* book), rather than facing extinction at the hands of more popular and captivating digital media, itself has become entangled with new media. In fact, the book owes some of its popularity not just to intersections with television (c.f. Oprah's famous book clubs) but also through complex digital systems for the distribution of physical texts. In the years since Striphos formulated this point, certainly Amazon.com's sophisticated digital systems for distributing but also recommending books have become even more complex. Still, historical development in the technology of textual media also does not *ipso facto* erase older ideologies that idealize certain uses of reading and certain ways of relating to books and authors. Collins (2010) asserts that while today's book cultures "may indeed rely on twenty-first century technologies," these cultures nevertheless "[draw] on early-nineteenth-century notions of reading as self-transformation, filtered through late twentieth-century discourses of self-actualization, all jet-propelled by state of the art forms of marking 'aesthetic experience'" (pp. 10). We must watch out for the ways that new technologies may sometimes paradoxically usher in *old* ways of reading and authorship.

In fact, "traditional" books likewise find synergistic connections with social media in what Simone Murray (2015) has recently called the "digital literary sphere," an umbrella term for the activity of readers as well as authors and organs of the publishing industry online. She notes that the emergence of this sphere has reorganized the relationship between readers and those latter two entities. On sites like Goodreads, readers themselves are often *critics* of literature, a status that has likely destabilized the hegemonic power of critics to treat readers as mere consumers to be advertised to. Indeed, in the context of online self-publishing, what counts as "literature" itself is now up-for-grabs. At the same time, she notes that authors and publishers increasingly turn to social networks as places where they themselves are active in order to

publicize and advertise particular works. Nakamura (2013) has observed that Goodreads is a space for literary criticism and conversation—"that most esteemed of discursive modes"—yet these very discussions are immediately co-opted and commodified by the network itself. The perspectives of Nakamura and Murray exemplify a tendency in media studies to consider the status of new forms of public, digitally-networked reading as labor (the fruits of which do not necessarily flow only to readers).

As Murray's analysis suggests, the "digital literary sphere" is not dominated by radically new types of books. Instead, "the vast majority of online literary discussion concerns traditionally linear, single-author narratives" (pp. 313). At the same time, literary scholars have also drawn attention to the new genres of texts that are emerging in the context of social media, such as new forms of serialized work produced for networks such as Twitter and Youtube (R. Page, 2013). Sorting out exactly the degree to which long-form genres and newer genres interact and co-exist remains an interesting question, one that this study will touch upon.

Given the complex and still-unfolding story of the book, it is not surprising that more optimistic theorists have pushed back against the notion that social media precludes "deep" reading. Dowling (2014) suggests that the expansion of book culture on social media, for instance, creates the opportunities for even deeper modes of literate attention than were previously afforded. For instance, considering some anecdotal data gathered from the book-centric online network LibraryThing, Dowling posits that the "complex movement back and forth between text, other texts, SMS with other readers" as well as other social media is simply a new kind of reading that is not any less "deep" than the old. While Sven Birkerts would vehemently disagree, Dowling is right to encourage a criticality about what exactly we mean when we say certain ways of reading are "deeper" than others.

### **Social Media, Literacy, and Literature**

Recent studies using various social scientific techniques have explored in more concrete ways the changing nature of the book, especially the rise of online readerly communities.

Using an interview-based ethnographic approach, Vlieghe, Page, and Rutten (2016) found that active users of a Flemish Facebook group for book lovers used the site for four main reasons: to "share reading experiences," to meet other readers, to forge identity (especially as a reader), and to encourage the participation of others on the site. They use the concept of the "affinity space" (Gee, 2005) to describe the social workings of this group. People were part of the site because they share a love of books, and so their involvement was designed to cultivate and support this "affinity" rather than to "maintain prior social bonds" (Vlieghe, Page, and Rutten, 2016, pp. 30). By way of Bourdieu (1984), this study also suggests the ways that online literary venues may function as spaces for displaying one's taste in books that makes plain one's "distinction." This study is an excellent example of the ways that integrating the voices of readers themselves can provide insights into the nature of the "digital literary sphere." At the same time, this study does not address what is *media specific* about this particular Facebook group; a traditional book club in a library or a home, after all, might fulfill many of the same functions.

Other studies have used quantitative techniques to study book-centric social networks. In their study of Goodreads, Thelwall and Kousha (2017) analyzed a large number of users in order to discover trends in the makeup of this site's users as well as their activities on the site. They found that two-thirds of Goodreads users were female, suggesting that participation on this site may be in some way a gendered practice; they were not able, however, to point to a reason for

this discrepancy. In fact, they found only minor differences in the ways that male and female users used the site. They also found that some users are more "social" (using features like *friending*) while others were more "book-focused" in their participation; in other words, the network facilitates different styles of participation. This study also revealed trends in the activity of users, with the users who have been members of the site the longest also the most active. In its conclusion, however, this study too does not really illuminate any media specific aspect of the network. Instead, the authors conclude that "Goodreads is an alternative to book clubs in contexts where book clubs are impractical or online interactions are preferable." Goodreads is a book club that simply happens to be online.

In another study, Thewall (2017) once again used quantitative (including, this time, linguistic) techniques to analyze gender on Goodreads in greater depth. Thewall found that there are clear gender preferences in literary taste on the site, with male and female users both generally preferring writing by authors of their own gender. Thewall also learned that there were differences in the genres preferred by male versus female users, with male users preferring genres such as the short story and female users preferring contemporary romance fiction. Using linguistic analysis, he also found that within their reviews Goodreads users focused on different aspects according to gender, with female users in some cases focusing on "relationships and romance." This fact suggests important gender dynamics occurring within the "digital literary sphere."

Digital networks also provide new spaces for the reproduction as well as contestation of literary canons. In a study that is intriguing in its creative blending of literary history and social scientific techniques (such as social network analysis), Finn (2012) considered the recommendation links between popular books on Amazon.com. These links between books were algorithmically generated according to histories of users' purchases; by mapping them manually (i.e. tracing which books were automatically linked to a book by following things to and from pages), Finn was able to shed light on users' literary values. Visualizing these links as networks of authors and texts, Finn demonstrated that the "canon" of Amazon in some way reproduces traditional canons, especially in its valorization of old works by white male authors like Mark Twain and F. Scott Fitzgerald. The surprising prevalence of the study aid CliffsNotes within these citation networks suggested the degree to which Amazon rankings are influenced by school-based syllabi and reading practices. At the same time, this network elevated some relatively recent fiction by writers of color, and so Amazon's algorithmic canons can themselves be "counterhegemonic."

Other studies have considered the ways that social media affords new types of literary authorship. Demonstrating the usefulness of an ethnographic approach, Vlieghe, Page, and Rutten (2016) studied an official "Twitter Fiction Festival" in which certain Twitter users were selected to produce literary texts to be circulated on this network. They found that, in this context, authorship was more like a "responsive performance" than a more traditional solitary form of authorship. Different Twitter-fiction authors navigated this new condition of authorship in different way. Above all, this study gives the impression of writers experimenting with the affordances of not just a new literary medium but a new relationship to an audience. Authors also experimented in their use of multimodal responses, including images and hyperlinks.

Vlieghe and his colleagues note a key limitation of their study: it was a one-time, official activity planned by Twitter itself, and so it would be useful to consider "bottom-up" approaches to the production of literary texts online. My recent work with Freedman and Hull (Booten, Freedman, & Hull, 2017) analyzes one such grassroots genre developed on Tumblr. This study



used quantitative as well as qualitative linguistic analysis to explore a corpora of "#ten word stories," a popular user-invented genre on this site. It found that many of these stories took the form of emotional *cri de coeur*, often using first and second person pronouns to create mini-tales of woe between an "I" (the lover) and a "you" (the beloved). We also found a strange ambiguity in terms of genre; a "#ten word story" might also be a "#ten word poem," for instance, suggesting that social media users may be both inventing new genres and blurring the boundaries between old ones.

In a larger analysis of the changing role of authorship on social media platforms, Myers (2016) notes the phenomenon of textual quotations from famous authors. He suggests that "authors merge entirely with other celebrities; the same Twitter feed could have the words of Maya Angelou and then Taylor Swift" (pp. 488). He analyzed several quotations to show the ways that quotations from literary works necessarily lose important context. However, since (as the author admits) the phenomenon of quotation emerged in the context of a somewhat different investigation, this study makes some claims about quotations that should be taken as hypotheses. For instance, it may be the case that one "does not actually need to read the book" in order to quote it, but it remains unclear exactly how people go about quoting on Twitter. Yet it remains an open question the possible relationships between quoting and reading, a gap this study will address.

In a more methodical investigation of a related phenomenon, Rowberry (2016) analyzed a large corpus of public highlights made by users of Amazon's Kindle (an e-book). This study used corpus linguistic techniques to analyze linguistic features of these highlighted passages relative to other textual corpora. He found two main types of highlights. First, there are those that (in reference to the ancient technology) he calls "commonplaces," referring to value-laden words (e.g. "vanity" and "virtue") as well as modal verbs (e.g. "ought"). The others are those that reference a key moment in the narrative, though these seemed to be less important. Rowberry's study is one useful jumping-off point for the present study; in general, however, its insightful observations about the linguistic nature of highlights—which may indeed be a type of "quote"—are nonetheless limited in the sense that they are not themselves contextualized in an ethnographic description that could shed light on *why* people value such passages, what they hope to gain from publicly highlighting them on a digital device, and how this practice itself may (or may not) transform the sorts of literate attention devoted to books during the process of Kindle-based reading.

**Other quantitative studies of quotability.** Rowberry's (2016) study of Kindle annotations quantifies the linguistic characteristics of these brief passages from books. I briefly mention here others studies have also explored "quotability," though not explicitly from books. Danescu-Niculescu-Mizil, Cheng, Kleinberg, & Lee (2012) studied quotations from an online movie database, finding that these quotations tended to have certain features that make them "generalizable," including fewer third-person pronouns, more indefinite articles, and fewer past tense verbs relative to present tense verbs. They also found that quotations tended to be lexically distinctive (containing rarer words than a comparison corpus) but not syntactically distinctive. Bendersky & Smith (2012) used a variety of complex features to computationally distinguish between quotes and non-quotes, finding that *abstract words* (words that are similar to Rowberry's (2016) "commonplace" words, such as "charity") and *quantifiers* (like "none") were the most helpful in distinguishing quotes from non-quotes. Studying a quotation of quotations from Tumblr, though not ones from books in particular, Booten & Hearst (2016) used statistical

techniques to observe characteristic "discourse-level patterns" in multi-sentence quotations. This study will attempt to build on this research by contextualizing quantitative findings about quotations as a linguistic genre in a larger, ethnographically-attuned discussion of quoters and the purposes of quotation as a practice.

### "Geeking Out" Online

Research—some of it ethnographic, some of it quantitative—has begun to explore directly the "fate of the book" through social scientific means. However, this can be considered just one small sliver of a much larger body of research that has explored the ways that people—including (and often especially) young people—can develop as literate subjects through their participation in digital and especially social media (Ito et al., 2009, Hull & Stornaiuolo, 2010; Lankshear & Knobel, 2011).

Lam's (2000) case-study of the online practices of a high school-aged L2 student has in recent years become a classic example of this mode of investigation. This student, Almon, became interested in building his own website as well as forging friendships with others on a GeoCities page devoted to Japanese pop-culture. For Almon, these online interactions became a way to practice English, yet the type of English Almon practiced online--"the global English of adolescent pop culture," rife with slang and allusions to shows and celebrities--was far different than the "standard English taught in ESL classes" (pp. 475). Lam concludes that Almon was actually doing much more than practicing a new language online; he was also fashioning a new identity for himself as a learner, one that was intimately related to the global pop-culture fandom and the type of English it valued:

Whereas classroom English appeared to contribute to Almon's sense of exclusion or marginalization (his inability to speak like a native), which paradoxically contradicts the school's mandate to prepare students for the workplace and civic involvement, the English he controlled on the Internet enabled him to develop a sense of belonging and connectedness to a global English-speaking community. (pp. 476)

Lam's case-study observes just one person in one community, but it gestures toward the way that digital technologies can give people access to new communities bound together by literacy practices. While this dissertation is not a study of language learning, Lam's study is a helpful point of reference in several respects. First, it is a reminder that young people especially engage in literacy practices that are often in juxtaposition to official, sanctioned literacy practices. Second, one cannot judge the value of a certain literacy practice simply by looking at it; it would likely be hard to glimpse the power of "adolescent pop culture" without hearing Almon's own story about his interactions with digital media. Third, the purposes of online literacy practices can be multiple and intertwined; the fact that this online world serves Almon as a language learner is inextricable from the ways that it serves him as a person in need of "connectedness."

Some of the digitally-enabled literacy practices that have been the most thoroughly studied are the various ways of engaging in online *fandoms*, including (but certainly not limited to) writing fan fiction (Black, 2009; Hellekson & Busse, 2006, 2006). Black (2007) explored in depth the ways that language learners in fanfiction communities provided direct one-on-one feedback as a piece of fanfiction's "beta" reader. Online fanfiction communities have evolved their own customs for providing critical feedback to authors, which Black divides into four main types: "OMG standard" (adulation meant to encourage the writer), "gentle critique," "focused critique," and "editorialized gossip" (in which beta readers discuss the fan fiction's characters as if they were real people, demonstrating a shared "high level of commitment" to the story). Black

also observed the etiquette underlying this practice, including a general injunction against the "flaming" that has characterized other, less self-consciously supportive online cultures. In this and other studies (see Black, 2005; Karpovich, 2006; Stein, 2006) a picture of fan fiction emerges: it is a way of reading texts—*deeply* reading them—through rewriting them in the context of a sophisticated community of peers.

In her recent work, De Kosnik's (2016) coins the term "rogue archives" to describe online archives that are largely maintained by everyday internet users and that contain the types of cultural production that have traditionally been excluded from the archives of mainstream institutions, including fan fiction. While De Kosnik notes that it is well-known that certain types of fandom are highly gendered activities, with most participants identifying as female, her research further shows that those who maintain fan archives are not just mostly female-identified but also often queer-identified. For many of these rogue archivists, the maintenance of fan archives takes on a political import that is also personal. Fan archives provide "safe spaces" for female and queer expression while also "[constituting] attempts to prove to the future that particular queer and female ways of being and making existed" (Kindle Locations 480-481).

Tumblr is a space for the circulation of fan fiction (Hillman, Procyk, & Neustaedter, 2014). As De Kosnik observes, it plays a particular role in relationship to official fandom archives, often serving as a place to "advertise" fan fiction. Moving forward with this study, it will be useful to consider the ways that Tumblr's rich culture of quotation is (or is not) a "rogue archive" in its own right.

### **Social Media and Audience**

The rise of social media means that individuals and the texts they produce circulate publicly in a way that poses new challenges.

Marwick and boyd's (2013) study of Twitter users suggests the complex ways that users imagine their *audience* in a new media space. Some of the users interviewed were "micro-celebrities"; for them, participation on the network can be a matter of building an audience--or even a "personal brand." Pursuing "micro-celebrity" status on Twitter brings about a tension between catering to the broadest possible audience (or perhaps pandering to the "lowest-common denominator" while still seeming like a real, "authentic" person). In other words, popular Twitter users must manage themselves as a brand without seeming to be managing themselves as a brand. In general, Twitter users described complex strategies of identity performance, such as carefully and self-consciously mixing in "personal" tweets with those that were more "professional" or simply self-censoring. Since a Twitter user's followers may be made of many different groups, including friends, family, and work associates, users must produce tweets of interest to these various groups while keeping in mind "nightmare readers" (e.g. bosses or parents) who might be offended by certain comments or personal information. While certainly identity-management occurs in offline spaces, it is uniquely challenging in social networks that "flatten multiple audiences into one." Social media users must often deal with this "context collapse."

Still, it is important to keep in mind that Marwick and boyd's study looked at only one particular network, Twitter. Audience and identity management on Twitter does not seem to be the same as it is on Tumblr. According to Shorey's (2015) ethnographic study of college-age female-identified Tumblr users, this network provides these users a safe space for expression. On Tumblr the widespread use of pseudonyms ensures a "culture of anonymity," one that in turn provides these young female users the freedom to express "culturally devalued emotions such as

sadness and anger" (pp. iii). In contradistinction to Facebook, in which users often described feeling lured into discussion or even debate, Tumblr offers them a space to express themselves in the company of like-minded peers without needing to explain themselves. This view of Tumblr's specific affordances as a network is a very useful starting-point for my own investigation, as I assume that the ways that readers pay attention to texts via social media may relate to their broader reasons for giving as well as seeking attention on specific networks.

### **Gaps In the Literature**

The research surveyed above has variously, seen as a whole, suggested that practices that have historically depended upon the book and print media, such as literature, are undergoing shifts both dramatic and subtle. Print media, however, have not vanished, and neither has the book totally evolved into some new post-book media form. Instead, these studies place readers and responders in a period of transition, synergy, and even symbiosis, as print literacies, bookish practices, old genres, and indeed print materials themselves become entangled with new media. Of particular importance is the way that print media and bookish cultures live on in the context of social media. Online book groups are the most obvious ways that social media supports rather than vitiates culture surrounding books. The practices of fan communities provide further evidence of the entanglement between old and new media on social networks.

My study of quotation culture seeks to address several gaps in this literature. First, we as yet know little about how the convergence of print and digital media has impacted how people pay attention to books; are "hyper" or "deep" attention at odds or complementary? While linguistic research, especially of the corpus or computational variety, has shed light on some of the linguistic characteristics of quotations, such studies suffer from a lack of ethnographic validity; to answer questions about attention, I suggest that it is crucial to study the nature of *texts* that circulate on social networks in the context of the *practices* that give these texts meaning. Finally, while sites such as Goodreads present themselves as social networks about books and for book-lovers, I suggest that focusing too closely on these obviously book-centric sites ignores the possibility that bookish culture occurs on social media more generally.

### **Theoretical (and Historical) Framing**

Already I have discussed the importance of considering *attention* in the study of literacy practices. Below I describe further theoretical vocabulary and perspectives that I found to be of use in carrying out the study.

### **Sharable Media**

The rise of digital technology has restructured the relationship between *producers* and *consumers* of media. While traditional organs of culture, such as film and television studios and publishing houses, are still significant, their heyday has passed. We are living in the era of "participatory culture" (Jenkins, 2012), in which members of the public are not passive "consumers of pre-constructed messages" but rather are actively involved in "shaping, sharing, reframing, and remixing media content in ways which might not have been previously imagined" (Jenkins, Ford, & Green, 2013, pp. 2). This wide availability of hardware and software for creating content, such as video, image, and sound editing equipment, in combination with the rise of the internet and, later, specific social media platforms, such as YouTube and Facebook, have broken down the traditional boundaries between those who produce media and those who consume it.

While a successful song can still be a "hit," a movie a "blockbuster," a book a "bestseller," a play a "critical darling," the rise of participatory culture has given rise to new ways for describing the relative success or failure of a text: whether or not it goes viral on one or several social networks. As Jenkins, Ford, and Green (2013) argue, metaphors of virality construe people as passive hosts to media that "infects" them and spreads through them. For this reason, they offer the term "spreadable" as an alternate term which is meant to draw attention to the agency that consumers have in *intentionally* sharing media. Likewise, they counsel against using terms like "stickiness" to describe the ways that texts capture the attention of consumers. "Stickiness," they argue, is a word that embodies the logic of advertisers who carefully design online media to capture users' "eyeballs" (i.e. their attention) which can in turn be sold to advertisers. To make a text "sticky," companies may target individuals through "user profiling," as is the case with algorithmically-tailored advertisements. "Spreadable" texts, on the other hand, are "those that various audiences may circulate for different purposes" (pp. 6).

Which term is to be preferred, "viral" or "spreadable"? While I agree that it is important to consider the agency of individuals and communities online, I begin with a more agnostic view. If thinking in terms of "spreadability" draws attention to *user* agency, thinking in terms of "virality" draws attention to *textual* agency: if a video gets shared millions of times, then certainly it has some qualities that have caused this spreading. Also, it is worth asking questions about the sort of agency that users have on social networks. When users click "share," "like," or "retweet," how much agency are they exercising, and what does this agency look like? Also, it may be the case that the regimes of "stickiness" and "spreadability" are more intertwined than they may first appear. On algorithmically-driven social networks such as Facebook, for example, there is a feedback loop between the two effects: the News Feed algorithm curates a user's news feed in part according to which posts have been widely shared by members of that user's social network (which may in turn lead to that user "liking" or "sharing" the post, furthering its propagation). For Jenkins and his colleagues, creating and sharing content on social networks is akin to *gifting*, which explains why companies that try to explicitly commoditize user-generated content often face a strong backlash. A more cynical perspective suggests that user-generated content always amounts to exploitation, since, without individuals who create and share content, companies such as Twitter and Facebook would be digital wastelands devoid of value (Scholz, 2008).

Whatever term is preferred, the fact that an individual social media user's texts can now reach a vast audience is inherently related to the concept of attention. For a text to go viral/become spreadable, it must capture the attention of large numbers of people. But it need not capture their attention *deeply*. Wasik (2012) describes our contemporary media landscape as one dominated by "nanostories," texts that quickly gather millions of views, likes, or retweets, only to just as quickly be forgotten. Often short and quickly consumable, these are not artifacts worthy of careful preservation; rather they are disposable (but addicting) trifles, cerebral snacks that offer little more momentary mental or affective stimulation.

Wasik's concept of the nanostory echoes a more common term used to discuss the new forms of contemporary viral media: the *meme*. Dawkins (2016) originally used this term in the way of suggesting that minimal units of culture thrive or vanish according to evolutionary pressure, much like the minimal genetic unit, the gene. Now, in the context of internet culture, the meme refers not to any single text but to a collection of texts, often humorous, that self-consciously refer to and remix each other (Shifman, 2013). Shifman notes that, unlike viral texts, memes are fairly conservative in the sense that they fit into identifiable genres, such as "image

macros" in which an image is overlaid with a brief text, a sub-genre of which is the "LOLCat," in which an image of a cat is overlaid with orthographically unorthodox words representing the thoughts of the pictured feline. These genres are not implicit; community members of online "meme-hubs" like Reddit discuss and debate these genres and their appropriate uses, while sites like knowyourmeme.com attempt to archive memes, capturing their often short life-spans in micro-historical descriptions. In other words, if a viral video is an *event*, a meme is a *practice*. As Shifman also observes, memes and one-off viral texts are both often meant to be consumed quickly and easily, since "when people understand something quickly and intuitively they are happy to forward it to others" (pp. 94). The concepts of the *meme* and *virality*, though distinct, both describe the new sorts of texts that have come to dominate an online world, texts that are finely-tuned to capture the fleeting attention of large numbers of people. Part of the effort of this study will be to understand the ways that quotation culture conforms to or resists the cultural logic of social media that privileges short and easy-to-digest texts.

### **Appropriation**

This study will also look directly at quotations that circulate on social networks. Quotations are interesting texts because they imply at least two authors: (a) the person who wrote the material that is quoted, be it a speech or a song or a line from a movie, and (b) the person who went through the trouble of repeating this language, quoting it. While this fact may seem simple enough, it draws attention to one of the reasons why quoting is a particularly interesting literacy practice: it requires one person to *appropriate* the words of another.

For Bakhtin (2010), appropriation is not just something that happens when we intentionally quote somebody. Rather, appropriation is the most basic precondition of acquiring and using language at all. All language is first encountered "in other people's contexts, serving other people's intentions" (pp. 294). To master any sociolect—the language of the academy, for instance, or the language of business—is not merely to parrot the words of academics or businessmen but to use these words for one's own "semantic and expressive intention" (pp. 293). For Bakhtin, the stakes of appropriating the discourse of others are always greater than mere "language acquisition." Since ways of speaking come bundled with ways of understanding, through navigating a world filled with competing sociolects and choosing which languages we want to appropriate, we by the same action determine who we are and what ideological positions we take in the world: "consciousness awakens to independent ideological life precisely in a world of alien discourses surrounding it" (pp. 345). This process, which Bakhtin calls "ideological becoming," is not accomplished without a great deal of struggle.

Quotation can be thought of as a particularly literal example of appropriation (see Knoeller, 1998). While it is possible to repeat verbatim the words of another, it is impossible to repeat the intentions behind these words (Bakhtin, 2010). What Bakhtin says is true of an individual word could also be applied to a quotation:

We have not yet learned from it all that it might tell us; we can take it into new contexts, attach it to new material, put it in a new situation in order to wrest new answers from it, new insights into its meaning, and even wrest from it new words of its *own*. (pp. 346-7)

Every instance of even the most literal quotation is a creative act in which the speaker must inflect the words of another with new intentions. At the same time, Bakhtin also cautions that it is possible to use the words of others in ways that abuse the original speaker's intentions, in effect taking their words "out of context." Parodic or otherwise uncharitable quotation can make "even the most serious utterance comical" (pp. 340). When this happens, one does not really

appropriate the words but rather treats them as "brute materiality" (pp. 340).

Thinking about quotation on social networks through the lens of Bakhtin's theorization of linguistic appropriation helps illustrate the stakes of an activity that may seem at first blush to be yet another way of wasting time on the internet. This dissertation explores the ways that the people who communicate with the words of others can do so in ways that go far beyond mere repetition: it can involve a great deal of care and creativity. At the same time, my study will be sensitive to the ways that quotation can itself produce or evince fundamental misunderstanding of the quoted text. Despite these dangers, however, Bakhtin's theorization of appropriation suggests the possible significance of online quotation culture in that it represents a practice through which people "try on" the words of others as their own.

### **Different Views on Practice**

My description of ethnography—and indeed the title of this dissertation—has made reference to the word "practice." In using this term I refer to the definition of "literacy practice" provided by Street (2001). A practice is in its most straightforward sense is a "pattern of activity," something people do over and over again (pp. 21). As I see it, when many people again and again post quotations to social networks such as Tumblr, the sum total of these events qualifies as a practice in this sense. However, as Street argues, there is also another, implied sense carried by the word "practice." More than mere repetition of events or the rules governing those events, practice points to something more latent, less immediately observable—those "broader cultural conception[s]" that govern how people both "think about" and "do" writing in specific "cultural contexts" (Street, 2001, p. 22). Street's definition is somewhat reminiscent of Bourdieu's (2005) notion of the *habitus*, a term that points to the ways that an individual's styles of being—ways of talking, walking, dressing, and reading—are embodied and naturalized ways of being that are distributed in ways that reflect the boundaries of wealth, status, and power within society at large.

In *The Practice of Everyday Life*, de Certeau (1984) distinguishes between different mechanism within practice. They are *strategy* and *tactics*:

I call a *strategy* the calculation (or manipulation) of power relationships that becomes possible as soon as a subject with will and power (a business, an army, a city, a scientific institution) can be isolated. (pp. 35)

The quintessential *strategists* are the city planners who, from high above, determine the flow of traffic; the business owners who determine what the workers will do and at what time they will do it; and the writer whose text builds a narrow path of meaning for the reader to follow. In contrast, the *tactic* is the process that "takes advantage of 'opportunities'" as they emerge in real time (since they, like guerrilla fighters, are without a space to call their own). The quintessential *tacticians* are the walkers whose light-footed, improvisatory wanderings defy any map or grid, the workers who "borrow" company equipment to surreptitiously complete their own projects, and readers. In fact, De Certeau's description of the nature of readers is worth repeating, as it continues to shape how we understand literary "consumption" versus "production":

Far from being writers—founders of their own place, theirs of the peasants of earlier ages now working on the soil of language, diggers of wells and builders of houses—readers are travelers; they move across lands belonging to someone else, like nomads poaching their way across fields they did not write....Writing accumulates, stocks up, resists time by establishment of a place and multiplies its production through the expansionism of reproduction. Reaching takes no measure against the erosion of time...it does not keep

what it acquires, or it does so poorly... (pp. 174)

This view of practice has been influential especially in fan studies, most famous in Jenkins' suggestion that fans often take hold of and rework texts toward their own purposes (Jenkins, 2012). Jenkins' study also notes a fundamental mismatch between De Certeau and new cultures of digital practices in which consumption and production increasingly blur: the traces of "consumers," such as their fan fiction, are *extremely* visible. Certainly, in the form of quotations and annotations, today's digital readers enjoy the "establishment of a place"—including those "rogue archives" (De Kosnik, 2016) already mentioned, those digital places that serve to "accumulate" and to "resist time."

Perhaps most binaries break if you push on them hard enough. Even if De Certeau's too-firm distinction between producers and consumers, between strategists and tacticians, may in some ways seem an ill fit for the intersection of production and consumption of digital media, it does not mean that his vocabulary is not useful. Instead, I suggest that it is interesting to notice places where everyday social media users operate in ways that are variously tactical *or* strategic. Furthermore, it may be interesting to consider where different types of strategies come up against each other—for instance, a social network's designers' plans for a network's use and those innovated by its users. Finally, simply because social media users in theory have a "space" to operate does not mean that some elements of their practice are not still space-less/tactical in nature.

### **Commonplace Books**

Viewing new technologies in light of old ones can help draw attention to unexpected continuities—an antidote to a knee-jerk assumption of the novelty of so-called "new media." I am not the first to suggest the similarities between digital quoting on Tumblr and one of the most well-studied ancient literacy practices, the keeping of a commonplace book. The general similarities are hard to miss. In a commonplace book, as on a Tumblr blog, readers keep quotations drawn from their readings.

However, general similarities can also be misleading. Within the literature surveyed in this chapter, Rowberry (2016) found it useful to consider the ways that highlighting passages on an Amazon Kindle is similar to commonplacing. In fact, his central claim—that the abstract terms (like "virtue") are themselves "commonplaces"—is interesting but also curious given that the term "common place" refers to organizational headings or "places" ("loci" in Latin, "topoi" in transliterated Greek) according to which readers would actively *categorize* their snippets of reading. To say that highlighting a text is more or less like commonplacing may illuminate some aspects of both practices while occluding other aspects.

In reality, the historical practice we now call "commonplacing" was not any single activity; likewise, the "commonplace book" is a term we use to describe a frustratingly diverse set of technologies, one that has evolved during its long history. The basic technology—a book for reading notes—took on different purposes during different periods of history. (The technology indeed does not begin with the term "commonplace book," which was a much later invention.) Scholars trace the notion of such an organization tool back to Aristotle, who first described "topics" of philosophical argument. Aristotle himself as well as other classical thinkers, especially Cicero and Quintilian, expanded these "places" and further theorized their relationship not just to logical argument but also to the arts of rhetoric. But it was Seneca the Younger who famously described the way that readers keep reading excerpts in separate categories the way a bee stores pollen (and where the bee converts it to honey). The staid



medieval *florilegia*, however, was filled not with examples of classical rhetoric but rather of austere logic and of moralistic Christian wisdom. The Renaissance "commonplace books," by contrast, were potent manifestations of that age's re-embrace of rhetoric and classical culture more generally. At different points, innovators reimagined the commonplace book by inventing new methods of organization. Even within a single historical moment, however, individual readers produced commonplace books that varied widely in content and purpose.<sup>1</sup>

There is obviously a danger in drawing comparisons between old and new media. In fact, perhaps some degree of generalization, even conceptual fuzziness, is required in staging a comparison (perhaps even a metaphorical relationship) between any new media and any old media as diverse as commonplacing. Still, I suggest that it is helpful to wonder not *how are contemporary textual practices like commonplacing* but *which version or versions of commonplacing most resemble, in their material manifestations and purposes, contemporary textual practices?*

### Research Questions

What kind of attention is at work in the practice of quoting from print media and print-based discourses? Hyper, deep, or some other kind yet unnamed? This is the underlying question that motivated this study, and it is the question that weaves together its three findings chapters, each of which explores a different array of data using different methods. The first of these explores Tumblr's broader quotation culture in order to make sense of what roles books and traditionally print-based discourses (e.g. literature and philosophy) play within it. The second views quotation from books on Tumblr through the words of those who engage in this practice. The third uses quantitative techniques to characterize quotations from Tumblr, especially those from books, and it does so with an eye towards the previous chapter's discussion of the purposes of quotation and how those purposes relate to these issues of attention. These three findings' chapters correspond to the following three research questions.

1. What role do print media, especially books, and the discourses traditionally manifested in these media play in the context of Tumblr's quotation culture?
2. How do quoters understand the value of quotes and the purpose of quotations that involve old, print-based media, and how do they understand quotation as fitting into their larger experiences of literacy?
3. What linguistic features characterize quotations on Tumblr, especially those from print-based media, and how to these features relate to users' purposes for quoting?

In answering the first question, I notice that the traces of books and bookish discourse on Tumblr do indeed suggest that practices of "deep" attention are connected to uses of social media but also that quoting involves other modes of attention, not strictly "deep" attention. In answering the second question, I document the way that quoters use quoting as a way of paying attention to books in ways that conform to and also complicate what the term "deep" attention typically means. Furthermore, I explore the ways that quotation is a key part of practices of self-expression and self-care, fundamental purposes of literacy that represent another way of experiencing texts "deeply." In answering the third question, I provide a typology of different

---

<sup>1</sup>This summary is yet another rehashing of a typical genealogy of the commonplace, made with close reference to both Havens (2001) and Allan (2010).

types of quotations in order to suggest the ways that different sorts of quotations from books solicit different forms of attention; by considering what types of quotations become popular on the site, I also consider what kinds of quotations solicit attention more "successfully."

This study ends with a conclusion chapter that synthesizes key findings while also recontextualizing them for two different disciplinary perspectives, literacy education and the history of media, both of which have a stake in understanding how digital technology is or is not transforming literate attention.

## CHAPTER 2

### METHODOLOGICAL INTRODUCTION: COMPUTATIONAL TEXTUAL ETHNOGRAPHY

#### Why Tumblr?

This study explores quotations and how people use them on Tumblr, a social network and a blogging platform. On the site, users can create their own blogs and "follow" other blogs.<sup>2</sup> Tumblr allows users to post one of eight different kinds of posts: "photo," "text," "quote," "audio," "video," "chat," "question," and "link." The photo-type emphasizes an image, such as a .jpeg or a .gif. The text-type allows for unstructured text input. The quote-type allows for the structured input of text. The video-type is for video files, the audio-type for sound files. The chat-type and question-type both enable conversation in different ways. The link-type is for hyperlinks. Central to the design of Tumblr is "reblogging," in which posts from one user's blog may be repurposed, showing up within another user's blog.

Figure 1 shows the main view of Tumblr from within the site. Media produced by users that one is following (as well as some recommended by Tumblr's algorithms) show up sequentially in one's "feed." The feed is inherently hard to capture in a single screenshot because it is effectively endless, but the user sees only a few pieces of media at a time. In Figure 1, one can see one post in its entirety, a quote-type post; scrolling downward would reveal the rest of the post consisting of a beige line drawing, a photo-type post. To the left of each post is the "avatar" (an image) of the user who posted or reblogged the post. From the feed it is possible to see how many notes ("likes" and "reblogs") a post has gotten; if the user who posted the quote added hashtags, these are visible here too. In the bottom right of each post are a series of buttons that allow different forms of interacting with the quote: sending it to another user, reblogging it, adding a comment to it, or "liking" it. On the right of the figure, blurred, are the avatars of users with whom one has recently chatted. At the top of the figure, to the right of the search bar, are various icons that trigger functionalities such as chatting with another user and creating a post.

Using the feed, it is possible to consume texts from a wide variety of users, though these same texts may be found by visiting each user's individual blog, where one may also "like" and "reblog" them. Individuals' blogs are highly customizable, both in terms of how they look and how users may interact with them. For instance, on one's blog one may include a tool that allows other users to submit posts for publication.

---

<sup>2</sup>In which case, these other blogs posts show up in the following user's "feed."

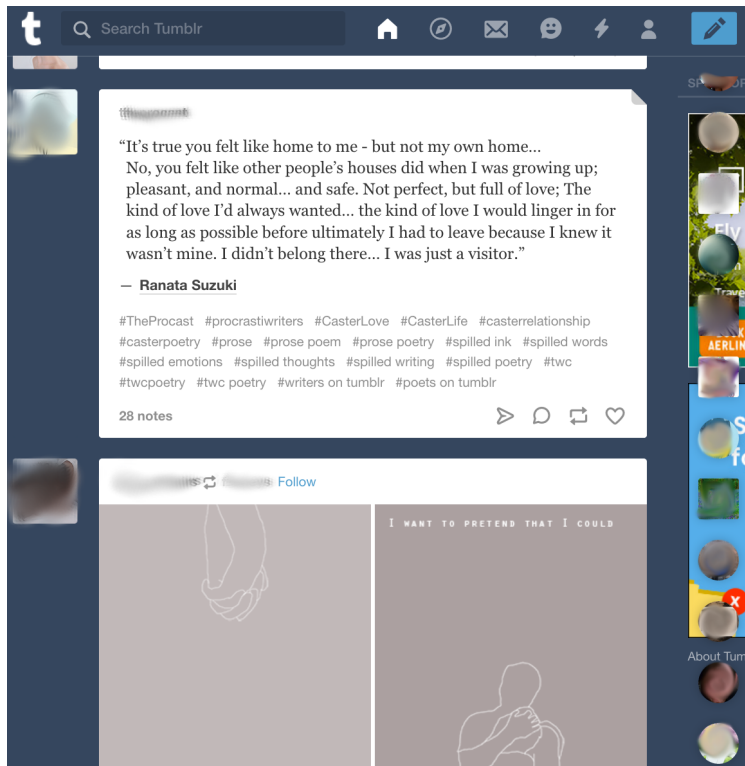


Figure 1. Tumblr's main "feed" interface.

I found Tumblr to be an interesting site for the study of digital quotation and related issues of attention for a variety of reasons. First, Tumblr is the only one of the major social networks (including Facebook, Pinterest, Instagram, and Twitter) to give its users the option of creating a quote-type post (Figure 2)<sup>3</sup>, though, to be clear, such posts represent only a small fraction of the overall Tumblr ecosystem (about 2.27% according to Chang et al. [2014]). Although quotes on Tumblr may take the form of other post types (such as the image-based photo-type), and although quotation is only one of the many activities that takes place on this network,<sup>4</sup> Tumblr is unique in that it actually was designed with quotation in mind. Next, since I am interested generally in education, especially of young adults, I chose Tumblr because it is a youth-dominated social network. According to a recent Pew study, 20% of internet users in the 18-29 age group user Tumblr, compared with 11% of the 30-49 group, 5% of the 50-64 group, and 2% of the 65+ group (Pew Research Center, 2015). This is a strikingly younger distribution compared to, for instance, Facebook; according to Pew, 48% of the oldest age group use this site (Greenwood, Perrin, & Duggan, 2015). Tumblr is a good place to look for the changing ways that young people read.<sup>5</sup>

<sup>3</sup> This quote-type post gives users the option to add the text of the quote, the source of the quote, and various hashtags. At the bottom right corner are icons for sending this post to another user, commenting on it, reblogging it or "liking" a post. The bottom left displays the total number of "notes," which consist of both "likes" and "reblogs." When one reblogs a quote, the originally hashtags do not automatically follow the post.

<sup>4</sup> Especially the photo-type, which consists primarily of an image file (such as a .jpeg or a .gif).

<sup>5</sup> Studying Tumblr does offer practical advantages over other sites. On Pinterest, quotations are also very popular; however, its lack of an API (Application Program Interface) precludes the sort

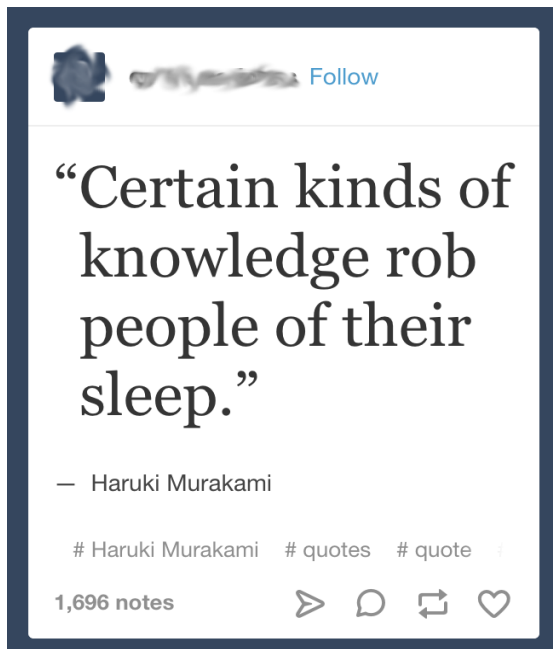


Figure 2. A quote-type post.

While an early-adopter of Facebook and a casual user of (or at least "lurker" on) Pinterest, Twitter, and Instagram, I had before this study very little experience with Tumblr. What experience I had had, however, suggested to me that Tumblr was somehow more creative and thus more interesting than other social networks with which I was familiar. I had mostly heard about Tumblr not through logging onto the main site (a portal through which one may search for blogs, posts, and users, as well as create and manage one's own blogs and accounts). Instead I had stumbled upon several individual blogs that had enjoyed at least a moment of popularity online. What struck me about these blogs was how self-consciously "niche" they seemed. On Tumblr, blogs devoted to Charlie Brown comics remixed with Morrissey lyrics<sup>6</sup> or photos of white men wearing the ill-fated Google Glass<sup>7</sup> seemed to be not minor oddities but mainstream attractions. I had yet to learn that this creativity (bordering on willful oddness) was in fact baked into the site's design and culture, especially through the relative anonymity the site provides its users. (It is not mandatory to list one's legal name on one's Tumblr blog, though some users do.) And I did not know at this point the degree to which book culture was intertwined with quirky and exuberant Tumblr fandoms (Hillman, Procyk, & Neustaedter, 2014).

I should also note that I chose this platform over social media platforms specifically-designed to facilitate the discussion of books, and even the sharing of quotations, such as

---

of large-scale data-analysis completed here. (An API allows for the programmatic gathering of data; without it, one is left to manually gathering or "scraping" data.) Tumblr's "chat" feature, in which users can by default message any other user, is much less restrictive than other networks' messaging features; Twitter, for instance, requires User A to "follow" User B before User B can send a direct message. Choosing another site would have meant using different methods. On Tumblr, one can send a message to all users by default, though some may opt to refuse messages from those whom they do not follow.

<sup>6</sup><http://thischarmingcharlie.tumblr.com/>

<sup>7</sup><http://whitemenwearinggoogleglass.tumblr.com/>

Goodreads. While I think that such platforms deserve more scholarly attention, I was particularly interested in seeking out "bookish" culture on Tumblr exactly because it is not "designed" for it, and where this culture (if it were to exist) would only be one community among many others. In the case of Goodreads, the designers have made many decisions to facilitate a certain way of interacting with literature. Studying Tumblr, a social network that might be called "content-agnostic," seemed interesting because it offered a chance to see whether or not, and how, users can transform a site to facilitate their own ideas about how to interact with literary texts. And, while my study focuses fairly narrowly on the sharing of bookish and academic quotations on this site, I at times was able to observe the ways that this practice bleeds into a few of the myriad other practices on the site.

### **From "Virtual Ethnography" to "Computational Textual Ethnography"**

It is still not uncommon to hear the word "literate" used as a synonym for "educated" or even "schooled," a linguistic fact that suggests the degree to which the prestige granted to certain forms of alphabetic literacy is deeply intertwined with a valorization of formal education. On the most general level, perhaps the most important intervention in studies of literacy in the mid-to-late 20th century was the movement toward considering the ways that "literacies" are as plural as the spaces and communities in which reading and writing (and other semiotic practices not purely linguistic) find purpose. At the heart of this long and enduring "turn," ethnographic methods proved to be indispensable, serving as the foundation of key works that illuminated what literacies look like as they are woven into the lives of real people as they navigate not only school but also leisure, work, and family-life (e.g., Heath, 1983, Brandt, 2001). This study—while not an ethnography in the most classical sense of this term—is motivated by an ethnographic desire to understand how one particular literacy practice, the circulation of quotations, functions in the lives of social media users, with particular focus on whether, when, and how quoters employ deep or hyper forms of attention.

Because interactions on Tumblr do not occur in physical space but rather are "virtual," I first looked to the emerging tradition of "virtual ethnography" (Hine, 2000). In the classical sense, ethnography is done in the field, a physical space that is unfamiliar to the anthropologist (be it a town, a school, or a scientific laboratory). Hine (2000) uses the term "virtual ethnography" to describe a new wave of ethnographies that are virtual in the sense that they study digital ("virtual") culture and that they are only "virtually" ethnographies and "not quite the real thing in methodologically purist terms" (pp. 65). Online worlds and cultures simply cannot be inhabited and studied with traditional ethnographic methods. Virtual ethnography is "adaptive ethnography" in that it "sets out to suit itself to the conditions in which it finds itself" (pp. 65). Old methods may be impractical, and new ones may need to be invented.

"Virtual ethnography" is ethnography carried out in virtual worlds, but it is also a sort of *meta* ethnography; it turns its gaze towards digital technologies and our commonly-held assumptions about them. Popular discourse abounds with "common sense" theories of what the internet is. (Social media has made everybody a narcissist. Twitter is a cesspool where any conversation about politics leads quickly to vitriolic hate-speech. Facebook is going to go extinct because young people think it is uncool.) Without outright denying the partial truth of any such understandings, virtual ethnography tries to generate a productive "puzzlement" by remaining open to ways of understanding the internet and its purposes that run counter to them. (This includes the researcher's assumptions as well as broader cultural ones.) In practical terms, my study is structured around certain research questions and indeed certain hypotheses about

quotation culture, inevitably grounded in my and others' preconceived notions of the connection between the internet and literacy, such as Hayles' (2007) distinction between "deep" and "hyper" attention. Yet, in order for this study to be ethnographic, I have tried to remain sensitive to the ways that my research questions, like any research questions, make assumptions that may run counter to those held by participants of the world I am studying.

There is something especially challenging, however, about studying *reading* through virtual ethnography. While traditional ethnographies of literacy have explored the complex ways that reading can be performed as an embodied act (Boyarin, 1989; Sterponi, 2007), observing reading "in the flesh" is particularly hard to do when the readers in question appear to the researcher only through the mediating device of the computer. Indeed, my research participants are quoters, and we were never co-present in physical space. Yet, for Hine, resisting common sense understandings about the internet also means rejecting any implicit attempt to position virtual communities against "IRL"<sup>8</sup> communities, instead exploring the ways that online and offline practices intertwine with each other. Though I was not able to see Tumblr quoters surfing the site, let alone in their everyday lives, I was able to explore the dimensions of their quotational practice via questionnaires and interviews.

While I believe Hine's insights into the conditions of computer-mediated research are invaluable, I also found that I needed to move beyond Hine's definition of virtual ethnography in certain respects. While strongly informed by virtual ethnography as formulated by Hine, my study is best described as an example of something slightly different, what I am calling "computational textual ethnography."

"Virtual ethnography," as its name suggests, was most concretely formulated as a way to study "virtual" spaces, such as gameworlds in which characters move and interact in realtime and three-dimensional space via avatars. Such worlds are not totally unlike traditional ethnographic field sites, only "virtualized" by computers. In them, one may enact a fairly traditional sort of participant observation. Tumblr is not a "virtual" facsimile of any offline world. It is a collection of texts—quotes, photographs, videos, and other kinds—woven together by screens, algorithms, and protocols. Thus I call this ethnography a *textual* one first as a reminder of this fact that a study of the texts that make up this network must remain front and center.<sup>9</sup> From my perspective, the textual nature of digital networks like Tumblr demands that ethnography be combined with discourse analysis. As Gee (2014) admits, "discourse analysis" is a broad and somewhat polysemous term.<sup>10</sup> My approach to discourse analysis is one that is adaptive and flexible; since

---

<sup>8</sup>"IRL" stands for "In real life." It is a common web-expression.

<sup>9</sup> Hine wonders aloud whether simply downloading and studying online, user-produced archives is enough to count as a form ethnography. While she ultimately advises a more engaged, participatory form of research, it is clear that such passive, archive-based methods may be *part* of an ethnography that "adapts" to the new forms of data readily available online. My study employs such archive-based methods, complementing them with more interactive methods designed to get at participants' understandings of Tumblr and quoting. Still, texts play a larger role in my study than they would in a standard "virtual ethnography."

<sup>10</sup>It can refer individually or simultaneously to (a) studies of language in use (sometimes called "pragmatics") and (b) studies of language above the sentence level, the complete structures and meanings of real-world texts. Some approaches that could be called discourse analysis are descriptive in nature, such as the text linguistic work of De Beaugrande and Dressler (1980) and Hoey (1983). Others, which go by the name of "critical discourse analysis," are political in

quotations are words but sometimes feature images, I at times blended discourse analysis rooted in linguistics with "multimodal" discourse analysis (Kress & Van Leeuwen, 1996). At the same time, I argue that performing discourse analysis on new kinds of texts often means inventing new analytic techniques, or at least remaking old ones to better suit new phenomena. In this respect I take inspiration from a body of discourse analytic research that continues to reckon with the media-specific characteristics of web-native texts (Androutsopoulos, 2011; Barton, 2015; Tagg, 2012; Zhou, 2009)<sup>11</sup>.

While I use questionnaires and interviews, I reject the notion that talking to somebody or observing them are the only ways of learning about their practices. Discourse analysis can be carried out as a part of a broader ethnographic approach of online literacy practices (Androutsopoulos, 2008), and I argue that this is particularly true in the context of a study of reading practices. An underline placed in a book, for instance, tells the analyst something (however vague) about what the reader found remarkable. Whether this marking was done in hasty highlighting or accompanied by an ornate manicure reveals still more. What is true about a physical book is even more true in the case of online quotations from books, since they are such obvious signs of readerly intention (and as well as attention). Reading itself is even more "visible" on social networks, since the quotations that readers circulate represent a very large, public collection of readerly traces. Thus discourse analysis, the analysis of texts, is a key part of understanding the practices through which social media users created those texts.

Quotations themselves are the most important texts that my study considers, yet by their very nature they are parts of other texts. Within the media ecosystem of Tumblr, digital quotations are texts that shift between contexts; they occur in specific blogs and move to other ones, for instance. Blogs, as contexts, are themselves another kind of text made up of smaller texts. And quotations may also be accompanied with what Genette (1997) calls "paratexts," especially hashtags. "Paratexts," according to Genette, are "vestibules" or "thresholds" that invite readers into text while also trying to ensure a "better reception" or a "more pertinent reading" (pp. 262). Studying quotations' paratexts as well as contexts can shed light on how quoters understand quotations and how they hope for them to be received.

My use of the word "computational" implies a shift not so much in perspective as methodology. As Ducheneaut, Yee, and Belotti argue (2010), virtual ethnography—while increasingly popular and accepted—is not always well suited to vast and complex online networks and archives. Especially salient is their critique that virtual ethnographies may not successfully describe an online world or practice if they focus only on a small sample of users from a very large collection of users. Their solution is to blend traditional ethnographic techniques (such as interview) with forms of computational data processing and analysis, not traditionally part of the ethnographer's toolkit. These include computationally-assisted quantitative analysis and visualization, for example. In fact, researchers who would not likely consider themselves ethnographers have imagined novel and powerful ways to study online social phenomena, such as how internet users shape their content production to specific

---

nature, with the goal of uncovering power relations as they constitute themselves in text (Fairclough, 2013; Wodak & Meyer, 2009).

<sup>11</sup>Some of these examples simply refer to the analysis of new types of texts, such as hashtags. However, there may also be a sense in which Discourse Analysis of web-based texts demands new analytic tools. Androutsopoulos (2011), for instance, has argued that the Bakhtinian notion of "heteroglossia" is fundamental to understanding multi-authored, web-based texts.



communities (Lakkaraju, McAuley, & Leskovec, 2013) or take up these communities' linguistic norms (Danescu-Niculescu-Mizil, West, Jurafsky, Leskovec, & Potts, 2013). For these studies to be ethnographic, however, they would need to contextualize large-scale quantitative findings with a "thick description" (Geertz, 1994) based in part on participants' understandings of their online worlds and practices. When I say that this dissertation is a *computational* ethnography, I mean that that it places into conversation "small" data, such as individuals' reflections on their own practices of quoting, with "big" data, such as aggregated collections of quotations but also hashtags and even usernames (computationally-collected data which I in turn analyze with computational techniques). In this dissertation, I planned some computational investigations from the beginning, but others emerged during the process of research, often stimulated by close inspection of particular data or through interactions with quoters. Quantitative and qualitative research unfolded in cycles.

### **On APIs, Hashtags, and Computation**

In Chapters 3, 4, and 5, I will discuss some of the more complicated ways that my research depended upon computational techniques, especially computational linguistic techniques. But when I say that this study is "computational virtual ethnography," I intend for the first word of this ungainly appellation to signal that I used computation in ways that are at once more mundane and more fundamental than computational analysis.

One of the ways that studying social networks is different than studying other textual corpora (such as microfiche collections of old newspapers) is that many social networks, Tumblr included, make available what is called an Application Programs Interface (API). In general terms, a social network API is a system that allows a programmer to write programs that interact with the social network in some way. Often APIs allow programmers to systematically harvest data from social network websites. Tumblr's API allows programmers to request original posts (not reblogs) that have been tagged with a certain hashtag; thus the hashtag is a fundamental *computational* unit. One can also request posts by a certain user. The API returns various pieces of metadata about a post, including when it was created, the user who created it, the complete list of its other hashtags, and the number of notes ("likes" or "reblogs") it received, and the type (be it quote-type, photo-type, etc.), and a link to the original post. The advantage of using the Tumblr API as a research tool is that it allowed for the mass collection of data that would simply not be feasible in any other way. (One can search for data on the site, but the results are relatively limited.) A key step in the research process was getting to know the API and to imagine what kind of questions could be asked and answered with the data made available by it. For instance, since the data returned by the API contained the type of the post, I was able to focus variously on quote-type posts or photo-type posts.

Picking focal hashtags was a key exploratory step in the research process. I explored tags by using the Tumblr search interface in order to see what kind of posts tended to carry various tag and by downloading posts via the API to simply see the quantity of data returned. My early impression of the "#quote" tag was that in general people use this tag to reference, as one may expect, quotations of all kinds, attributed to various authors. Within the broader culture of quotation on Tumblr, to examine issues of attention, I studied the specifically "bookish" tags: "#book quote"/"#book quotes." These were tags that focused on quotes from books specifically; these tags especially are important to my study because they directly represent the intersection of new/digital and old/print media. "#Literary quote" was another candidate focal tag, but after downloaded sample data via the Tumblr API I noticed that this tag seemed to be

used much less frequently than "#book quote" and "#book quotes," as the API returned much less data.

Hashtags, as research tools, are much like nets; they do not catch everything (a user need not tag a quote from a book "#book quote"), and, because users can use the same tag to mean different things, some unwanted texts<sup>12</sup> may be caught. The computational ethnographic approach is useful because it triangulates different types and scales of data to confirm findings, using the strengths of "big data" and "small data" approaches to make up for some of the weaknesses of each.

### **On Analytic Units**

Different scholarly disciplines resort to different epistemologies, different understandings of the nature of truth and knowledge. One of the hidden challenges of doing "interdisciplinary" work such as this study is making use of and reconciling distinct epistemologies and indeed different customs of interpreting data.

As computational research, this study often depended upon the quantification of certain units of data that can be precisely defined—for instance, the frequency of a word, a hashtag, or a grammatical construction. While there may have been some creativity involved in the generation of an analytic unit, the actual analysis is not subjective; whether a quote contains the word "love" or not was determined by an algorithmic process for counting words, not by me. In such cases, the algorithm intervened between myself and the data, analyzing it "objectively."

On the other hand, as ethnographic research, this study depended upon qualitative analysis and interpretation of the data, the first step of which was often "coding" or breaking up data into conceptual units. In the context of interpretive qualitative research, such as this ethnographic work, coding "solo" is a valid strategy (Saldaña, 2015). As Saldaña (2015) notes, coding is both a rigorous social scientific technique and an "art" that depends upon the researcher's unique creative perspective as well as unique vocabulary for describing the world. The description of qualitative analysis provided by Elliot and Timulak (2005) resonates with my experience as a researcher:

Creation of categories is an interpretive process on the part of the researcher in which the researcher is trying to respect the data and use category labels close to the original language of participants. On the other hand, ideas for categories also come in part from the researcher's knowledge of previous theorising and findings in other studies.

Categorising is thus an interactive process in which priority is given to the data but understanding is inevitably facilitated by previous understanding. It is a kind of dialogue with the data. (Elliot and Timulak 2005, pp. 154)

No doubt my knowledge of "previous theorizing" colored my interpretation from my data; this is not a fault but rather a condition of the production of new knowledge (i.e., new interpretations). As some of these codes are complex and make subtle interpretive distinctions but also connections between phenomena (for instance, between, the various interlocking reasons quoters give for quoting), my application of them depended upon my deep, ethnographic familiarity with the data in the process of solo-coding. Coding was thus not merely a process of applying labels to data but a complex working-through of the underlying logics of practice that these labels represent; this complexity itself is manifest most clearly in the writing of findings in which codes

---

<sup>12</sup> For instance, in Chapter 3 I discuss the ways that "#book quotes" may also be quotations *about* books.

are a mere starting point, structuring a deeper "thick description" (Geertz, 1994) of the data. Furthermore, in presenting data in the various findings chapters I have aimed to present ample data corresponding to various codes; thus readers may get a sense of the diversity and complexity of data I have described with these codes.

The term that I have used to describe this research, "computational textual ethnography," thus marks two different ways of parsing reality—*computationally* (i.e. quantitatively) and *ethnographically* (i.e. qualitatively)—which the study uses in tandem.

### **On Words**

One of the threads that links together the three findings chapters is the word "deep," as in "deep attention" and "deep reading." Its counterpart "hyper" is another thread. I have used these words so far as if their meanings are at once durable and transparent. Part of the work of *writing* the dissertation (in contradistinction to the work of "conducting research") was to struggle to use these abstractions to describe real practices and real people, to notice what they obscure as well as reveal, and to find out what else they may mean.

### **What "Books"?**

To posit simply that this dissertation explores the connections between "the book" and "digital media" is true but also misleading in its simplicity. A "book" might include anything from a paperback novel bought in an airport to a scholarly collection of folktales (originally maintained through an oral tradition) to a medieval medical manuscript.

Part of the goal of this dissertation was to take this term—"book"—and, using it as a torch, noticing what textual objects glint under its light. (It is not much of a spoiler to note that medieval medical manuscripts are not widely quoted on Tumblr, nor are quotations from telephone books or automobile owner's manuals.) At the same time, to probe issues of deep and hyper attention, I did have a sense of what broad sorts of texts I wanted to investigate specifically. The first was *literature*—a term I use to include anything from popular science fiction to drama to poetry. To be clear, even these categories are not necessarily equivalent to "the book"; poetry, after all, may be seen in periodicals or recited from memory. The second broad category was *philosophy*. Philosophy too can exist without the book, though the strong connection between philosophy as an academic discipline and the proliferation of manuscripts in university settings dates back to the late Middle Ages (Schmitt, Skinner, & Kessler, 1988). I was particularly interested in philosophical quotations because of the highly *academic* (and otherwise abstruse) nature of much philosophical discourse.

As the reader will see, certain genres of possibly "bookish" quotations revealed themselves to be more crucial to quotation culture, and to certain communities within this broader culture, than others. At various points in the dissertation I focused on whichever genre or combination of genres helped me to answer my research questions most thoroughly.

### **Phases of Research**

The next three chapters of this dissertation (3, 4, and 5) report the findings from three distinct phases of research. Each phase makes use of a different subset of methods under the overarching banner of "computational textual ethnography" and provides answers to one of the three major research questions for the study, in the process shedding light on the matter of *attention*. Included in each findings chapter is a discussion of the methods specific to answering the research question for the chapter.

## CHAPTER 3

### THE GHOSTS OF OLD MEDIA ON TUMBLR

Phase 1 focuses on *digital texts*, observing the complicated ways that old and new media converge in the form of quotations on Tumblr. It answers the question: What role does print media, especially books, and the discourses traditionally manifested in these media play in the context of Tumblr's quotation practice? My goal was to understand whether the apparent hyper attention implied by quoting was or was not connected to deep attention to the books from where the quotes originated. I thus looked for (a) places where the "hyper" attention of quotations seemed to connect with literacy practices involving "deep" attention (especially the presence of books and bookish discourse in quoted form) and (b) ways in which some forms of quoting may themselves seem more "deeply attentive" to their source material than others.

Before I could answer my first research question, I had to understand Tumblr as an archive so that I could focus on those aspects of the Tumblr ecology that would be most relevant. I gathered data that could be immediately found on the site, which included quotes that were posted on various blogs, the sources of these quotes, the tags that users applied to these quotes, and the pictures they used to illustrate them. In general, the methodologies applied in this chapter all fall under the rubric of "discourse analysis," including close analysis of particular texts, noticing trends in larger corpora of texts, and quantifying discourse features. Following my overall methodology of "computational textual ethnography," this chapter operates at different scales, using quantitative techniques to produce a "high-level" view of certain discursive phenomenon on the site while also examining particularly important examples (*sans* computational assistance). The view is not so much a generic "bird's eye view" but what might be called a *hawk's* eye view, as it interrupts its high-altitude gyres to swoop down, snatching a piece of data for focused analysis. In this Phase, I strive to appropriate and invent methods that are well-suited to the new sorts of texts—including "contexts" and "paratexts"—that surround quotations on Tumblr. By combining quantitative and qualitative discourse analysis, I was able to account for the importance of books (and book-based reading) in quotation practice in general and the ways that quoting entails other forms of "paying attention" to texts that do not always fit neatly into the "deep"/"hyper" binary.

The analyses I performed during this phase reveal some of the ways that the ghosts of dead or dying literacies haunt Tumblr. On this site, old media—especially the book—as well as the genres that have traditionally depended upon old media can often be found. As is the case with most ghosts, however, the passage to the "afterlife" has left these old media and their attendant genres transformed. And, as with most ghost stories, this one comes with a warning.

Those who read ahead in search of confirmation that the rise of social media spells the decline of literacy itself will be disappointed. Yet so will those who expect to find a vision of forms of reading liberated from our hidebound traditions and thus perfected. A more truthful approach to the fate of old media on Tumblr is one that explores the multiple and at times contradictory ways in which books and historically bookish discourses are appropriated. This chapter is at its heart an account of some of these contradictions.

#### **Data and Methods**

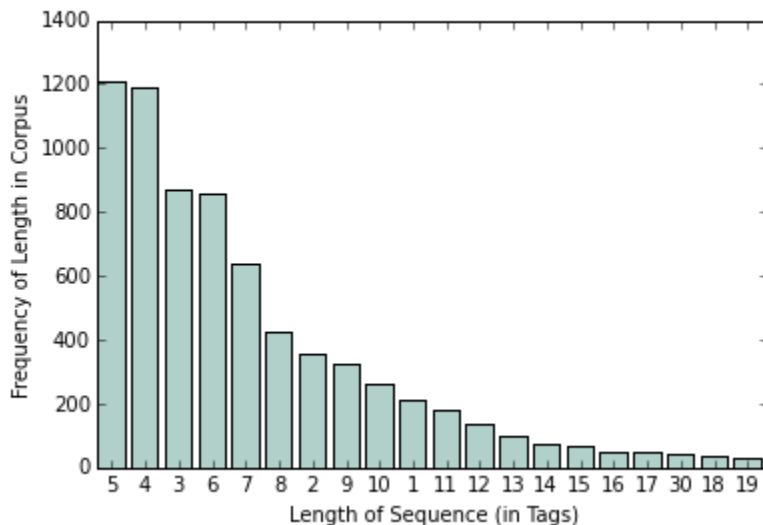
## Data

Here I describe in more detail the corpora of Tumblr data I analyzed for this phase. First, however, I describe what I call the "unfiltered" corpora, large downloads of data which I did not study directly but from which I created other corpora that I did analyze. The Unfiltered GQ Corpus included 74,908 quote-type posts that had been tagged "#quote" created between May of 2011 and April of 2015. The Unfiltered BQ Corpus included 32,868 quote-type posts that had been tagged "#book quote" or "#book quotes" representing the same timeframe. The Unfiltered BQ Photo Corpus included 8,680 different photo-type quotes tagged "book quote" created between Sep of 2009 and April of 2015.

**Hashtag corpora.** Hashtags, as key terms that quoters use to describe their quotations, offer a useful way of exploring quotation practice from an aggregated level. Looking at the distribution of these tags can also provide clues as to the way that quotation practice involves books as well as the sort of attention that users give to these books in quoted form. The GQ (General Quotes) Hashtag Corpus consists of each tag used by each of the 36,152 users from the Unfiltered GQ Corpus (n=273,370). (I only count a tag once for a user even though the user may have used it many times.) The BQ (Book Quotes) Hashtag Corpus consists of each tag used by each unique user in the BQ Corpus (n=103,256).

Hashtags can be thought of as individual units, yet Tumblr users usually apply several tags, not just one, to their posts. I define a "hashtag sequence" as a set of hashtags applied to one particular post. For instance, a post may be tagged "#quote," "#beautiful," and "#literature." These sequences are useful because they essentially represent the co-occurrence of semantic units. The GQ Hashtag Sequences Corpus consists of one tag sequence, chosen at random, for each user in the GQ Unfiltered Corpus. Figure 1 shows the 20 most frequent lengths of hashtag sequences in this corpus.

Figure 1: Top 20 Most Common Lengths of Tag Sequences, GQ Hashtag Sequences Corpus



**Sources Corpora.** Another way to analyze the ways that quotation practice does or does not involve quotations from books is to ascertain who are the most quoted people or entities. Analyzing sources too can point to the degree to which quotations draw upon books; in combination with analysis of hashtags, analyzing sources can shed light on whether or not books (in quoted form) do receive attention within the textual economy of Tumblr. Analyzing sources in particular can also suggest the types of books and authors that receive the most attention on

the site. However, while this information is necessary groundwork, it must be put into conversation with other information to suggest *how* quoters pay attention to these authors. One of the advantages of studying quote-type posts (versus, for instance, the photo-type) is that this type comes with a source field in which users are implicitly encouraged to enter the sources of their quotes. From the unfiltered GQ and BQ corpora, I created two corpora of sources. Sources were extracted and processed using basic text processing techniques, with each user contributing a source (such as "Taylor Swift") at most once to the corpus, no matter how many times she quoted this figure. Sources were processed according to simple text-processing techniques. For instance, source strings were split at certain punctuation marks as well as other dividing character sequences such as " by ", breaking a string such as "Harry Potter by J.K. Rowling" into "Harry Potter" and "J.K. Rowling." The GQ Sources Corpus represents 83,741 sources. The BQ Sources Corpus represents 39,281 sources.

**Simple BQ Corpus.** During preliminary analysis I became aware that some quotations tagged "#book quote" and "#book quotes" were *about* books or reading. Some of these seemed to be *from* books, others not. To understand whether books were really receiving attention on Tumblr, I needed to tackle the ambiguity of these tags. Later (in my discussion of Phase 3) I discuss more complicated corpora designed for linguistic analysis of the quotes themselves. For preliminary analysis, I created a "simple" corpus consisting of one quote from each user in the Unfiltered BQ Corpus (n=7,260), including its source string. My purpose for this corpus was also quite simple: to determine whether most book quotes indeed come *from* books (rather than are merely *about* books), a possible confusion that emerged in the course of this phase's analysis.

**Philosopher Quote Corpus/Hashtag Corpus.** Since philosophy is traditionally a discourse communicated through print texts (i.e. philosophical monographs such as *Thus Spoke Zarathustra*), philosophical quotations represent an important point of possible connection between quoting on Tumblr, print literacies, and academic literacies. I chose to explore how people apply tags specifically to philosophers in order to glimpse how readers understand them. I created a collection of seven philosophers<sup>13</sup>, ranging from those I deemed to be generally-known (e.g. Socrates) to the less-so (e.g. Angela Davis). I used these philosophers' names as hashtags in order to download all posts available via the API tagged, for instance, "#angela davis."<sup>14</sup> I took steps to further filter out quotes that were likely *about* rather than *attributed to a* philosopher.<sup>15</sup> The total number of these quotes, the Philosopher Quote Corpus, was 3,847. From these posts, I collected hashtag sequences as I did in the Hashtag Corpora above; these I call the Philosopher Hashtag Corpus. These hashtags represent signs of the specific ways that quoters pay attention to philosophical texts on Tumblr.

**BQ Photo Corpus.** Images are another important type of paratext that quoters use to

---

<sup>13</sup>Angela Davis, Camus, Emerson, Heidegger, Karl Marx, Nietzsche, and Socrates.

<sup>14</sup>In cases where the personage seemed relatively unambiguous, such as Heidegger, I searched only the last name—"#heidegger" and "#socrates," for instance. In other cases I searched for the entire name; for instance, I searched "#angela davis" and "karl marx," to make sure my search did not pick up any quotations by the Marx Brothers.

<sup>15</sup>I did this by filtering out any quote by Socrates, for instance, that contained the word "socrates."

frame the words of others. Analyzing the images that quoters use to illustrate quotations, like analyzing quoters' uses of hashtags, provides a glimpse into the ways that they understand specific authors as well as quotation itself. From the Unfiltered BQ Photo Corpus, I chose one post for each blogger and then randomly collected a small sample (n=200). As with hashtags, these photos offer signs of the ways that quoters pay attention to books through Tumblr quotes.

**Possible misquotes.** Understanding the degree to which misquotation occurs on Tumblr is an important way of ascertaining whether quoters' engagement with quotations is a "deep" one. From the website Wikiquotes I collected a list of commonly misattributed quotes (n=8). In picking these particular misquotations to investigate, I used my best judgment to focus on those that were somehow "bookish." (For instance, I chose a misquotation wrongly attributed to Voltaire—known as an author—not one misattributed to an American baseball manager.)

Hunting for specific misquotes does not give a sense of whether misquotations are popular. (If a misquotation happens on Tumblr and nobody reblogs it, does it really make a sound?) I chose to further search for misquotations by re-examining the Philosopher Quote Corpus; since many of the quotations on the Wikiquotes website were misattributed to philosophers, I hypothesized that philosophical quotations in particular may provide further evidence of misquotation. For the philosophers in Philosopher Quote Corpus, I also created ranked lists of the top 5 most popular quotes (in terms of likes and reblogs) that I had downloaded to see if any were misquotes.

**Usernames Corpus/Blog Description Corpus.** Unlike sites such as Facebook, where using one's legal name is the norm if not mandated, Tumblr allows and even encourages pseudonyms. Usernames provide an interesting example of the ways that Tumblr users represent themselves. I noticed that some users gave themselves names that could be considered "bibliophilic" such as *katerinareadyabooks* or *mybelovedlibrary*. (These are invented but characteristic examples.) Arguably, these are signs that one is so "deeply" invested in literature that it has become part of one's identity. To investigate this matter more thoroughly, all of the 7,260 individual usernames in the Unfiltered BQ Corpus were collected into a single corpus, the Usernames Corpus. I also noticed that on Tumblr users add personal information to a public profile, including age, gender, and interests. Using the Tumblr API, I downloaded 1,000 descriptions from users in the Usernames Corpus to make a Blog Descriptions Corpus.

**Audience Usernames Corpus.** When analyzing book quotes on Tumblr, I thought that it was important to explore not just who posts quotes (represented by the Usernames Corpus and the Blog Description Corpus) but also who consumes them by "reblogging" or "liking" them. I was curious as to whether consumers were less "bibliophilic" than those who produce them, which would signal that book quoters might bring book quotes to the attention of those who are not as bookish. The Tumblr API only offers limited information about book quoters; one may retrieve only the 50 most recent notes ("likes" or "reblogs") attached to a post. For each of the 7,260 blogs in the Unfiltered BQ Corpus, I randomly selected one quote from each; I then chose 1,000 that had at least 1 note ("like" or "reblog"). Using the API, I randomly selected one of the "reblogging" blogs, leaving 826 blog names. (This number is smaller than 1,000 because for some posts the sole "liking"/"reblogging" account was the original creator of the post, not a different user, and because the API sometimes simply fails to return the original post.) The Audience Usernames Corpus represents each unique name out of these 826 names, a total of 805

names.

**A focal quote.** During and between the various forms of analysis described above, I stumbled into a variety of quotes as well as blogs that "stood out" for one reason or another, and I made notes of these as I explored. To conclude this phase, I returned to one particular quote (attributed to Heidegger) from one particular blog. For reasons I describe fully in Chapter 5, it stood out to me as a "telling case" (Mitchell, 1984), not necessarily a "typical" one. As Mitchell notes, a telling case is one that "serve[s] to make previously obscure theoretical relationships suddenly apparent" (pp. 239). In particular, my ethnographic intuition suggested that this case might reveal something "previously obscure" about the presence of bookish, and specifically academic, discourses on Tumblr.

## Analysis

Here I describe the techniques I used to analyze the corpora described above. In analyzing the corpora, I appropriated and invented a number of techniques in order to shed light both on the role of books and bookish discourse within the context of quotation on Tumblr and on the ways that quoters attend to bookish quotes.

**Quantifying tags/visualizing tag relationships.** To explore the Hashtag corpora, I first simply tabulated the top 30 most common hashtags. I assumed that the distribution of hashtags would roughly follow a Zipfian distribution, as has been shown in case of Twitter hashtags (Cuhna et al., 2011); my choice of presenting the top 30 tags was arbitrary, as is common in the analysis of hashtags and similar data (Hermida, Lewis, & Zamith, 2014; Manikonda & Kambhampati, 2014; Saxton et al., 2015). This tabulation gave a basic sense of what key terms characterize quotations in general as well as the more specific type of quotations represented by the BQ and GQ Hashtag corpora. In particular I noticed whether top tags in the GQ Hashtag Corpus contained any references to books or reading, and whether top tags in the BQ Hashtag Corpus contained any tags that suggest that quoters quote in the process of "deep" reading.

However, merely counting tags does not provide any information about the *relationship* between tags. Tags, like words in general, are polysemous. What do they *generally* mean? A good way of answering this question was to explore the way that certain tags co-occurred with each other, following Firth's principle that one "shall know a word by the company it keeps" (Firth, 1957). To visualize the relationship between tags as they occur in hashtag sequences (GQ Hashtag Sequences Corpus), I first adopted an approach used elsewhere to describe the relationship between fan-fiction tags, also on Tumblr.<sup>16</sup> In general, this approach uses hierarchical clustering, an old and venerable method for exploring the underlying "taxonomies" within data (Gower & Ross, 1969; Johnson, 1967). Given a set of hashtag sequences, hierarchical clustering (and related visualization methods) can be used to produce a "tree" in which more closely-related tags (i.e. tags often used together in tag sequences) appear closer together and are linked closely by "branches" (as if they were similar species in an evolutionary taxonomy). On a technical level, the approach is to (a) create a matrix representing the Jaccard coefficient measuring the co-occurrence of any two tags, (b) cluster these tags using SciPy's (Jones, Oliphant, Peterson, & others, 2001) hierarchical clustering algorithms, and finally (c) visualize the hierarchical organization of clusters as a dendrogram.

---

<sup>16</sup><http://mycommonreader.github.io/Hierarchical-Clustering/>



I applied this technique to the top 100 most frequent tags in the GQ Hashtag Sequences Corpus. In less-technical terms, the point of this visualization is to take a list of tags ("#love", "#hope", "#depression", "#poetry" ...) and essentially re-order the list so that closely related tags are placed close together on a "tree." This visualization method helped me to make sense of the complicated folksonomic organization underlying hashtags in ways that merely looking at raw frequencies could not. As the computer-generated visualization is quite complicated and information-rich, I have manually annotated it in order to draw attention to key phenomena (and away from clusters of tags that, in the interest of clarity and brevity, I do not have occasion to mention). This visualization further helped me understand how books and bookish discourses fit into a complicated ecosystem of Tumblr quotations.

**Analyzing Sources Corpora.** For each of the Sources corpora, I simply tabulated a frequency distribution of sources, which could refer either to an author ("Haruki Murakami") or a work ("Norwegian Wood"). I tabulated the top 20 most quoted authors, replacing each work in the tabulation with its corresponding author. I omitted certain attributions that did not seem to actually refer to a source (for instance, "https," which often appears in source strings that include links). I analyzed these sources in terms of basic information about the sources' profession (Are they authors? Sports personalities? Movie stars?), national origin, gender, and race. This provides a sense of to what degree authors of books receive attention on Tumblr. As I discovered, sources on Tumblr quotes can contain additional information, such as hyperlinks to a web-based text from which a quote was extracted. I therefore was able to calculate the number of links to certain websites, such as Goodreads. Quotes taken from a website point to a way of quoting that is arguably less "deep" than quoting in the process of reading.

**Analyzing Simple BQ Corpus.** As noted earlier, I initially found that some quotes tagged "#book," "#book quote," and "#book quotes" occasionally seemed to be *about* books. Here I explore this ambiguity with more rigor, calculating the number of tags containing references to words having to do with reading: "book," "books," "reading," "read," "reader." I also examined the quotes containing these words to see if any of these quotations *about* books were also *from* books.

**Zooming in on philosophic tags.** For each collection of tags representing a specific philosopher in the Philosopher Hashtag Corpus, I created a frequency distribution of the most common tags. I then analyzed the top 15 tags for different philosophers and inductively coded them; in particular, I drew a distinction between those tags that suggest familiarity with the author's work, and even possibly an *academic* understanding of this work, and those that do not. These hashtags can suggest different ways of "paying attention" to traditionally bookish discourses on Tumblr, with some ways being arguably deeper than others.

**Analysis of photo-type posts.** My analysis of specific quotations began with an inductive coding of the Photo-Quote Corpus meant to capture recurring types of visual representations. Examples include photos of books and GIFs/stills from film/TV shows. During this process, however, I became intrigued by the ways that certain types of photographic quotes hinted at other communities and practices on Tumblr. In the end, I found it most illuminating to use my coding of the Photo-Quote Corpus as a starting point, after which I used the Tumblr

search interface to find related examples tagged with other hashtags that I discovered through the exploration of the Photo-Quote Corpus, such as "#book photography" and "#fanedit." By exploring these related tags, I was able to get a sense of how quoting connects to other practices within Tumblr; these practices themselves entail different ways of paying attention that do not necessarily fit into the "deep"/"hyper" binary.

**Checking for misquotes.** I understand misquotation to be clear evidence of a way of quoting from books that could not reasonably be called "deep." In analyzing the known-misquotes (from Wikiquote), I simply noted whether or not I was able to find them repeated on Tumblr. I also noted interesting instances in which the quote was attributed correctly, some of which signalled a deeper engagement with the source text. For the most popular quotations from the various philosophers, I individually used Google Books, Google Scholar, and Google's main search to try to locate the quotation's source text. In some cases, this led me to online discussions or even other scholarly papers about a specific misquotation. Still, it is impossible to prove a negative; absent a perfectly complete archive of *all* of Ralph Waldo Emerson's utterances (including diary entries, spoken conversations, etc.), one cannot prove that he did not in fact at some point utter the famous slogan "Just do it." My identifications of misquotes must be taken as the result of good-faith effort to attribute a quote to a particular author.

**Analyzing usernames/blog description corpus.** For each name in the names corpus, I programmatically checked whether it contained one of a few strings: "book," "read," "novel," and "liter" (as in "literary"). I then tabulated the percentage of names containing these strings. For each description on the Blog Description Corpus, I looked for individual words (after lowercasing and tokenizing the description): "read," "reading," "reader," "books," "book," "bibliophile," "bookish" or "bookworm." I take the presence of such words as a clue that the blog is about books, reading, or literature—in a word, that it is a "bibliophilic" blog whose creator is likely a reader who has a "deep" relationship to literature.

**Analyzing Audience Usernames Corpus.** To analyze this corpus of names, I noted how many of the names contained references to books or reading ("book," "read," "novel," and "liter"). I then compared these percentages to those for the Usernames Corpus and Blog Description Corpus using a chi-squared test. This comparison provided a sense of whether or not the audience of this particular quote were themselves "bibliophilic" (i.e. whether they generally seemed to "pay attention" to books to the same degree).

**Analyzing focal quote.** I analyzed the focal quote attributed to Heidegger by considering its "contexts," both the original context of the blog to which it was posted and the "secondary" contexts of the blogs to which it was reblogged. I engaged in a "thick description" (Geertz, 1994) of the blog's contents, analyzing the kinds of posts it contained (especially whether they were also philosophical or academic texts). To analyze the blogs that have reblogged the focal quote, I invented my own method that is meant to condense into a readable paragraph a fragmentary representation what these blogs are "about"—what I call "context collage." To create a collage, I simply clicked on the link to each blog and described this blog's most recent post, usually a sentence or less—i.e. "a GIF of a cat running in a circle" or "a black and white photo of the singer Lana del Ray." I compiled these brief, condensed descriptions into a single paragraph. I present this method as a form of "textualization"; a few words can be worth a

thousand pictures. I found this novel methodology helpful for exploring the ways that the accounts that reblogged the focal quote were or were not similar in nature to the blog that originally posted it.

## Findings

### **Old Media, Including Books, are an Important Part of Quotation Practice in General**

I begin with a quantitative exploration of the hashtags of quote-type posts tagged "#quote" (BQ Hashtag Corpus). This analysis shows the crucial role that books play within quotation practice in general.

Table 1 shows the top 30 tags out of a total of 66,823 unique tags; the median number of occurrences of a tag in the corpus is 4.09. The tags in this table are very diverse in nature; clearly, however, tags that have something to do with old media, including books, are among the most frequently used tags.

"#Book" and "#book" are both in the top 30 most popular tags (#12 and #21). "#Literature" (#20) and its abbreviated form, "#lit" (#22) are also popular. However, these tags are bested by some tags that are clearly about other sorts of media, specifically those related to music: "#music" (#7) and "#lyrics" (#9). Tags having to do with the book (a media form) and literature (a discourse that has traditionally circulated in books) are an important part of Tumblr.

I briefly note that certain thematic concerns are also popular. "#Love," for instance, marks just under 16% of the quotes. "#Life" too is popular, occurring on over 10% of quotes. These two tags in particular do seem to be of particular importance; the next three tags, "#personal," "#poetry," and "#lyrics," occur on 4.3%, 4.3%, and 3.9% of quotes respectively. Thus, while "#love" and "#life" are not an order of magnitude more popular, they do stand out. The fact that they do occur so frequently suggests that these tags may point to overriding thematic concerns that characterize quotes regardless of provenance (whether lyrics or books, for instance).

Figures 2A and 2B present a hierarchical cluster of the top 100 most frequent tags used in the #quote dataset, arranged according to their frequency of co-occurrence.<sup>17</sup> This visualization helps elucidate the meaning of some ambiguous quotations.

---

<sup>17</sup> As the visualization was quite larger, I have divided it in two, though there is some overlap between them; a red dot helps orient the viewer toward this overlap.

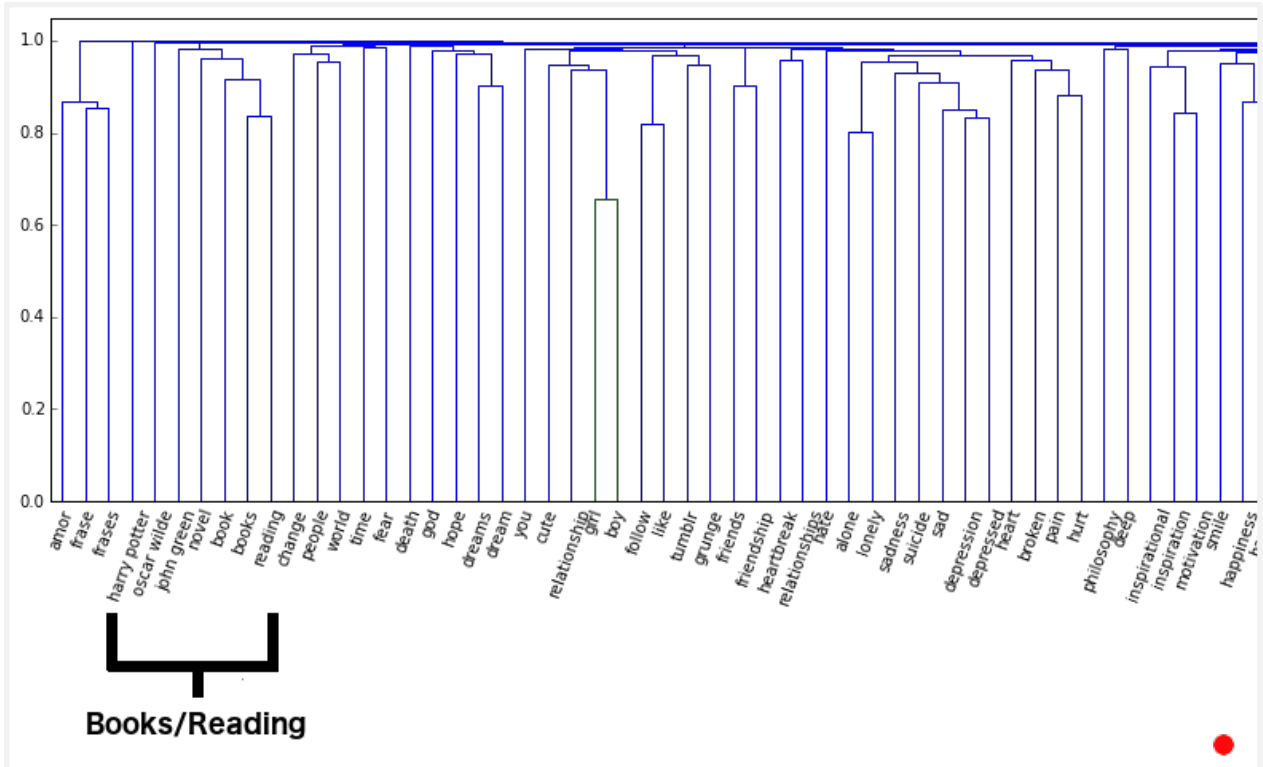


Figure 2A. Hierarchical clustering of top 100 hashtags used in the GQ Hashtag Corpus (LEFT).

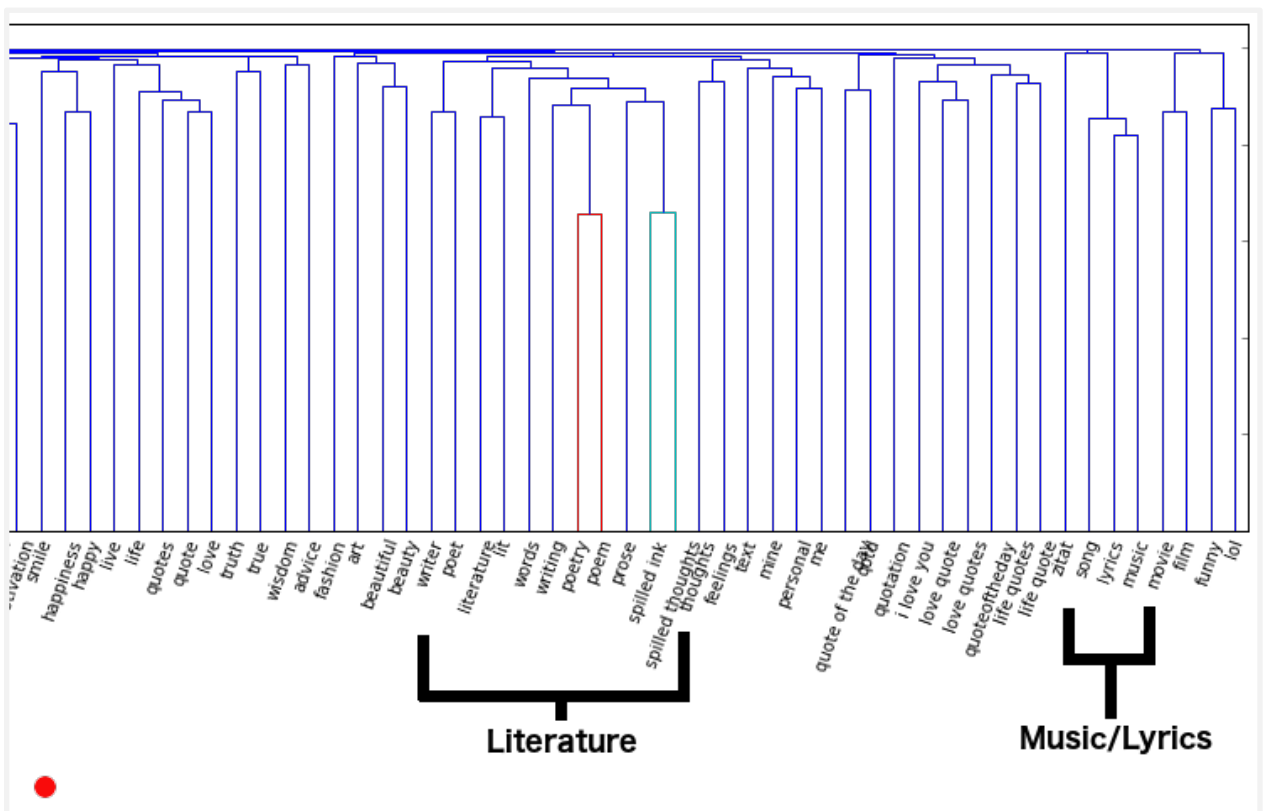


Figure 2B. Hierarchical clustering of top 100 hashtags used in the GQ Hashtag Corpus (RIGHT).

"#Book" and "#books" cluster with "#reading" as well as tags related to specific authors and books, including "#john green" and "#harry potter." It is worth spelling out that the tag "#reading," in combination with the tag "#book," suggests a relationship between quoting and a more traditional type of literacy: reading books. I interpret the fact that "#john green" and "#harry potter" cluster near "#book" as evidence that quoters often quote from John Green's fiction or the J.K. Rowling's *Harry Potter* novels. (It is noteworthy that *Harry Potter*, a franchise that includes not just bestselling series of novels but a chain of blockbuster movies; within Tumblr's quotation practice, it seems, this franchise is most strongly affiliated with the "#book" manifestations.)

As one would expect, "#music" and "#lyrics" cluster near each other, suggesting that they co-occur frequently. However, this clustering also reveals some of the unexpected complexities of quotation practice. Tags like "#literature" and "#lit" do not cluster with "#book" but rather nearby a series of other tags, all of which have something to do with textual production or writing: "#words" and "#writing," "#poetry" and "#prose," "#spilled ink" and "#spilled thoughts." These final two tags are different than the others in that they are part of Tumblr's particular idiolect, its specialized hashtag vocabulary. As one Tumblr blog explains<sup>18</sup>, "#spilled ink" is a tag that writers on Tumblr use to describe their own textual production. Often these take the form of short poems or stories, and for these writers "#quote" can mean not an excerpt from another work but a short work itself. The clustering of tags suggest that "#poetry" and "#prose" tend to refer to such user-authored "quotes."

"#Philosophy" is another tag which could possibly represent a connection between digital and traditional and even academic literacies<sup>19</sup>. Yet "#philosophy" does not cluster anywhere near "#book." Nor does it cluster with "#spilled ink." Rather it appears on its own, yoked together with the tag "#deep."<sup>20</sup> It is difficult from merely this visualization to sense exactly what constitutes philosophy on Tumblr, other than discourse that is somehow "deeper" than other discourses.

Table 1 thus adds further proof that old media is a large part of quotation practice, yet it also cautions against a simplistic understanding of this textual ecosystem. I begin this study with this "high level view" of Tumblr's quotation practice as mapped through some of the most common hashtags in order to suggest that this practice is diverse and complicated.

Another way to explore the relationship between digital quotation and old media is to examine the most popular sources for #quote. Table 2 lists the top 25 sources for quote-type posts tagged "#quote." As this table makes clear, books and bookishness may be an even more important part of quotation practice than the simple tabulation of hashtags suggested. Not one of the top sources hints at "#music" or "#lyrics." Rather they are mostly authors of one sort or another. Only Albert Einstein (#6), Marilyn Monroe (#18), and Audrey Hepburn (#90) could be said to be known not primarily for their writing. ("Unknown," "anonymous," and "me" also point to quotations that are not quoted from any particular media or, in the case of those attributed to

---

<sup>18</sup><http://spilling-your-ink.tumblr.com/spilledink>

<sup>19</sup>Since philosophical discourse often occurs through print material, including books (e.g. Kant's *Critique of Pure Reason*, Heidegger's *Being and Time*).

<sup>20</sup>Notice that on the y axis the two tags are joined together at a level immediately below 1.0. This suggests that the algorithm found a strong correlation between these two tags but that in turn they did not seem to be related strongly to any other clustering of tags.

"me," are simply invented.<sup>21</sup>)

The authors listed represent a diversity of genres. There is contemporary fiction (Chuck Palahniuk, Haruki Murakami, Paulo Coelho) and even Young Adult fiction (John Green). There is poetry (Sylvia Plath, Charles Bukowski) as well as philosophy (Friedrich Nietzsche, Albert Camus). Ralph Waldo Emerson might conceivably fit into either category<sup>22</sup> and indeed searching the site for his name revealed examples of both his poetry and prose appearing in quoted form. Given this diversity it is perhaps fitting that Oscar Wilde—the poet/dramatist/essayist—tops the list.

Still, this list in other ways is not particularly diverse. Of the authors listed here, all save Anais Nin, Sylvia Plath, and Maya Angelou are men. Most of the writers are American, British, or European.<sup>23</sup> Most of these authors originally write or wrote in English, and almost all are white. And most are dead.

Analysis of popular hashtags and their co-occurrence has shown that Tumblr quotation practice is connected to books in different ways. First, quotes themselves may draw from "#books" themselves, perhaps in the course of "#reading." In other words, quotation involves the book itself as a media, and reading as a way of interacting with this form. It is also clear from examining popular sources that quotation practices on Tumblr involve the quotation of authors from various genres that have historically circulated via books and print media more broadly, including literature but also philosophy and verse.

Table 1.

*Most common tags for posts tagged "#quote."*

<b>Rank</b>	<b>Tag</b>	<b>Percentage Users Who Use</b>
1	quote	100.0%
2	love	15.9%
3	quotes	11.9%
4	life	10.2%
5	personal	4.3%
6	poetry	4.3%
7	lyrics	3.9%
8	sad	3.6%
9	music	3.5%
10	words	3.2%
11	inspiration	3.0%
12	book	3.0%
13	writing	3.0%
14	mine	2.9%
15	poem	2.9%

<sup>21</sup>This phenomenon is similar to the "#spilled ink" quote; on Tumblr a quote may often be a pithy message (or even poem or story) invented by a blogger (Booten, Freedman, & Hull, 2017).

<sup>22</sup>The *Norton Anthology of American Literature* includes both his philosophical prose, such as "Nature," and his poems, such as "Brahma."

<sup>23</sup>Clear exceptions to "American, British, or European" include Haruki Murakami (Japan) and Paulo Coelho (Brazil). Anais Nin was born to Cuban parents in France. Albert Camus was born in French Algeria.

16	me	2.7%
17	text	2.5%
18	truth	2.3%
19	thoughts	2.1%
20	literature	2.1%
21	books	2.0%
22	lit	1.9%
23	depression	1.9%
24	song	1.9%
25	quoteoftheday	1.6%
26	art	1.6%
27	happiness	1.6%
28	true	1.4%
29	happy	1.4%
30	beautiful	1.4%

Table 2.  
*Most common sources for posts tagged "#quote."*

<b>Rank</b>	<b>Tag</b>	<b>Count</b>
1	unknown	542
2	me	481
3	oscar wilde	311
4	anonymous	267
5	john green	221
6	albert einstein	204
7	charles bukowski	161
8	sylvia plath	148
9	haruki murakami	144
10	mark twain	141
11	william shakespeare	136
12	f. scott fitzgerald	131
13	friedrich nietzsche	121
14	ralph waldo emerson	113
15	chuck palahniuk	113
16	ernest hemingway	113
17	paulo coelho	106
18	marilyn monroe	104
19	anais nin	100
20	kurt vonnegut	99
21	albert camus	97
22	neil gaiman	94
23	audrey hepburn	90
24	maya angelou	88
25	stephen king	83

### Book Quotes Are Different Than Quotes in General In Certain Ways (And Not In Others)

While books are a key part of quotation in general, the practice of the "#book quote"/"#book quotes" is in some ways different than the broader practice of quotation.

The most striking difference between book quotes<sup>24</sup> and quotes in general<sup>25</sup> can be seen by examining once again hashtags and most-quoted sources. Table 3 lists the top sources from quote-type posts tagged "#book quote" or "#book quotes." There is some overlap between this list and Table 2. John Green, Haruki Murakami, Neil Gaiman, F. Scott Fitzgerald, Paulo Coelho, Oscar Wilde, and Chuck Palahniuk occur in both lists. However, there are also important differences between the two lists. First, the most popular "#book quote" sources are more homogenous in terms of genre. While it is possible to find poets, philosophers, and other genres further down the list of top "#book quote" sources, in general this tag is much more strongly associated with fiction. To be more specific, its top sources lean more toward specifically Young Adult fiction, authors of which include Cassandra Clare, Rainbow Rowell, David Levithan, Stephen Chbosky, and Mark Zusak.

Table 3.

*Most common sources for posts tagged "#book quote" or "#book quotes."*

Rank	Tag	Count
1	john green	291
4	haruki murakami	122
6	neil gaiman	113
7	cassandra clare	108
8	stephen king	102
9	rainbow rowell	102
10	david levithan	102
11	f. scott fitzgerald	101
12	stephen chbosky	95
13	jane austen	92
14	paulo coelho	88
15	markus zusak	85
16	jodi picoult	84
17	oscar wilde	84
18	sarah dessen	83
19	lauren oliver	80
20	jonathan safran foer	79

Table 4 illustrates the top 30 hashtags applied to quote-type posts tagged "#book quote" or "#book quotes" out of 32,285 unique hashtags. An obvious difference between this list and Table 1 is that the former features many more words related to books and reading—tags such as "#reading," "#novel," and "#booklr." This last tag, another piece of Tumblr's idiolect, according to one Tumblr user, a tag used to mark "a loose collection of people with blogs who like

---

<sup>24</sup>What I call those quotes drawn from the "#book quote"/"#book quotes" tags.

<sup>25</sup>i.e. quotes tagged "#quote."



books."<sup>26</sup> This word (like "#bookworm") is a term not so much for describing the material of the quotation as for describing the quoter's identity and marking her affiliation within a bibliophilic community.

Table 4. Most common tags for posts tagged "#book quote" or "#book quotes."

Rank	Tag	Count
1	book quotes	67.4%
2	book quote	42.8%
3	quotes	34.2%
4	quote	29.3%
5	books	27.3%
6	book	17.9%
7	love	15.3%
8	life	8.7%
9	literature	7.7%
10	reading	7.5%
11	john green	5.0%
12	bookworm	5.0%
13	love quotes	4.8%
14	life quotes	3.8%
15	words	3.7%
16	lit	3.7%
17	writing	3.2%
18	booklr	3.1%
19	poetry	2.8%
20	novel	2.7%
21	sad	2.7%
22	death	2.6%
23	truth	2.5%
24	love quote	2.5%
25	depression	2.2%
26	inspiration	2.1%
27	currently reading	2.1%
28	looking for alaska	2.1%
29	happiness	2.0%
30	pain	1.9%

Tags can also help illuminate the ways in which quotation is linked to other forms of literacy. The reading related tags "#reading" (rank #10) and "#currently reading" (rank #27) are among the most frequent tags used with book quotes, though they are not among the most popular tags used with book quotes. Thus different tags suggest different practices of quoting; as one would expect, book quotes tend to come from books in the process of reading.

At the same time, comparing Table 1 and Table 4 reveals that certain tags are common in both. *Emotions* are important characteristics of quotations in both Table 1 and Table 4. In fact,

---

<sup>26</sup> According to the Tumblr blog bookoisieur (bookoisieur.tumblr.com).

from these lists it appears if those quotations having to do with negative affect—quotes that are "#sad", for instance—are as popular if not a bit more popular than those having to do with positive emotions, such as "#inspiration." "#Depression" is also popular in either; this tag, less popular than "#sad," suggests not just an affective state but a clinical diagnosis. As with quotes in general, however, "#love" stands out as a major thematic concern (15.3%) followed by those about "#life" (7.7%); the presence of these themes near the top of both Table 1 and Table 2 suggests that the major themes addressed by book quotes may be largely similar to those addressed by quotes in general.

### **Hashtags Suggest "Deep" ...and Less "Deep" ...Forms of Literate Attention**

That the tag "#currently reading" is one of the most popular tags applied to book quotes suggests the ways that the quotation can occur as part of a traditional literacy practice, reading. Looking at more idiosyncratic tags, it is possible to see the penumbra of readers' diverse reading practices.<sup>27</sup> A blogger notes that she is going to download a text by the quoted author onto her Kindle<sup>28</sup>. Another points out that the quote contains the words from the "whole book" which "stuck with" her the most. Yet another notes the fact that she is already halfway through the quoted book, having started it only earlier that day. Another confesses to rereading the quoted book, despite being short on time. Another admonishes other bloggers to read the quoted book, because if you haven't "you are living your life wrong." Others simply document the strong feelings evoked by reading: "#this book is already turning into a favorite holy cow". Echoing the ways that quotations also tend to deal in (often negative) emotions, other tags register the immediate heartbreak inflicted by a particularly poignant passage: "i keep reading this book and i keep crying." In these tags, it is possible to glimpse the ways that quotation and print literacy are intertwined—sometimes even temporally, with quoting occurring in the process of reading. In general, such tags are evidence of the fact that at least some quoters devote their "deep attention" to the books that they quote.

However, a "#book quote" is not necessarily evidence that a reader is in fact extracting a quote directly from a book in the process of reading. Out of a sample of 7,260 sources from the BQ Sources Corpus, 250 of them (less than 4%) contained a hyperlink. Some of these suggest other sources of quotations. For instance, 188 of these were hyperlinks linking to the website Goodreads. This site, a social network for book lovers, features pages devoted to user-submitted quotes from books. In such cases, a "#book quote" may indeed be *originally* from a book but discovered and reposted from an external website. In fact, the number of quotations produced this way is likely higher than the count of those that explicitly link to Goodreads, since a blogger could just as easily repost a quote from Goodreads (or another site) *without* acknowledging it.

**A note on the ambiguity of tags.** As mentioned in Chapter 2, some number of quotes with tags like "#book," "#book quote," and "#book quote" are ambiguous. In my analysis of the Simple BQ Corpus I found that book quotes (as one would expect) most often quotes from books. Yet 6% of "#book quote[s]" in this corpus contained one of the following words: "book," "books," "reading," "read," "reader." In this sense, they may be said to be "#book quotes" in that they are *about* books. Yet, as I alluded to in Chapter 2, it is not quite so simple. Some quotes of

---

<sup>27</sup>Bloggers often use Tumblr tags in order to add free-form comments to posts. In the case of quotations from books, often these take the form of comments about the process of reading.

<sup>28</sup>An e-book reading device sold by Amazon.

this type are attributed to non-book sources, yet many others are attributed to books. In the latter case, it is unclear whether these are "#book quotes" are called this because they are *from* or *about* books—or perhaps both. In any case, it is interesting to note that one of the functions of quotes from books on Tumblr is to share perspectives on books and reading themselves. Often these perspectives are positive, such as this one from Austen's *Pride and Prejudice*:

*I declare after all there is no enjoyment like reading! How much sooner one tires of any thing than of a book! -- When I have a house of my own, I shall be miserable if I have not an excellent library.*<sup>29</sup>

It seems plausible that one of the functions of book quotations is to celebrate reading itself—a sort of meta-commentary on reading itself carried out by quoters (who are also devoted readers).

### Quoters Use Academic Vocabulary

In general, quoters—even those who share book quotes—do not seem to quote as part of any formal, academic reading. This can be deduced from the relatively few mentions to school reading in the BQ Hashtag Corpus or the GQ Hashtag Corpus. Searching for quotes containing the words "class" and "school", it was possible to find tags that suggest that quoters quote from texts assigned in formal learning contexts:

idk i'm reading this for school and i really liked how he phrased this  
we were reading this book for school  
i have been reading it for my english class  
i'm starting to like lit class wow weird??

However, such tags that link quoting to school reading were very rare—5 in the case of BQ Hashtag Corpus, 15 in the case of the much larger GQ Hashtag Corpus. (As a comparison, this is an order of magnitude fewer than those quotes that link to Goodreads.) Yet quotation practice is not purely a space of leisure reading where readers voice their excitement about books. It is also a place for readers to share quotations from authors in a way that points to a disposition that is more academic than leisurely. This can be seen especially in the case of philosophical quotations.

Tables 5 through 11 present the frequency with which other tags are applied to these philosophical quotes in the Philosopher Hashtag Corpus. First, for five of these philosophers, one of the most popular tags refers to the title of one of the particular author's works.<sup>30</sup> What is also striking about these tables is the preponderance of words that deal with concepts that are clearly couched in the vocabulary of philosophical discourse. In the case of Angela Davis (Table 5), "-ism" words abound—"#racism," "#feminism," "#capitalism."<sup>31</sup> These terms make sense in light of the fact that Angela Davis is a black feminist philosopher whose works critique the modern prison system as a racist/capitalist institution that profits from the mass incarceration of people of color. That "#prison industrial complex" is mentioned so frequently suggests the degree to which quoters of her work understand this very specific critique of a very specific aspect of

---

<sup>29</sup> Throughout this study, I depart from traditional APA style in the way I re-present quotations from Tumblr and elsewhere. No matter the length of the quotation, I indent present it as a block quote and put it in italics.

<sup>30</sup> Angela Davis (*Are Prisons Obsolete*), Karl Marx (*Das Kapital*), Nietzsche (*Thus Spoke Zarathustra*), Camus (*The Stranger*, *The Myth of Sisyphus*), Emerson (*Nature*). This last one is admittedly ambiguous, as nature is a theme of much of Emerson's work in general.

<sup>31</sup> Words with the "ism" suffix are signs of academic discourse (see Höpfl, 1983 for a discussion of the rise of "-ism" in the history of Western thought).

contemporary culture. Karl Marx (Table 9) is often quoted with a similar rhetorical effect. Once again, quoters often use "-ism" words as tags: Marxism, "#communism," "#capitalism," and "#socialism." And, once again, these terms point to social and structural phenomena. Note that these lists are generally devoid of thematic tags that quoters use to characterize quotes in general, such as "#inspiration" and "#love."

It is interesting to compare the table for Karl Marx and Angela Davis with the table for Heidegger (Table 8). Heidegger's philosophy devotes relatively little attention to economic institutions and structures of power (at least in his most quoted text, *Being and Time*). Still, co-occurring tags suggest a concern with labeling Heidegger as certain type of philosopher—as working in the tradition of an other "-ism," "#existentialism"—and another philosophical movement, "#phenomenology." Such terms do not just label what Heidegger is "about" but place him into an academic history of philosophical movements. That "#dasein," Heidegger's famously challenging neologism, also appears in this list suggests that quoters of Heidegger are interested in the slippery vocabulary of academic philosophical discourse. Camus (Table 6) is connected with "#existentialism" and also an "#absurdism," while Emerson (Table 7) is connected with "#transcendentalism." Such "-isms" do not play a central role in the tagging of either Nietzsche (Table 10) or Socrates (Table 11), though this may have to do with the fact that the relationship between these individuals and any one philosophical movement is less obvious than in the case of other philosophers. Still, it is interesting to note that in the case of Nietzsche, Camus, and Socrates especially, vague thematic terms that are among the most popular tags applied to quotations in general are also frequently applied to them—"#life" in the case of Nietzsche, Socrates, Camus, and Emerson, and "#love" in the case of Nietzsche and Socrates. On Tumblr, an excerpt from Emerson's works may be considered both an example of "#transcendentalism" and of words that provide "#inspiration." The academic and the demotic freely commingle.

Table 5.  
*Tags on posts tagged "#angela davis."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
quote	39	614	100.00%
quotes	34	571	11.90%
racism	33	33	0.16%
feminism	22	23	0.43%
prison industrial complex	19	19	0.00%
politics	16	28	0.40%
prison	15	17	0.04%
are prisons obsolete	14	14	0.00%
revolution	12	20	0.05%
capitalism	12	29	0.07%
prisons	10	10	0.00%

black women	10	10	0.00%
change	8	23	0.79%
activism	8	8	0.05%

Table 6.  
*Tags on posts tagged "#camus."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
albert camus	297	298	0.20%
quote	138	614	100.00%
quotes	106	571	11.90%
philosophy	77	370	0.89%
existentialism	65	92	0.10%
life	56	211	10.21%
albert	40	40	0.02%
absurdism	39	39	0.00%
literature	39	103	2.07%
the stranger	38	38	0.02%
the myth of sisyphus	36	36	0.01%
love	33	96	15.87%
absurd	29	30	0.02%
citation	29	42	0.37%

Table 7.  
*Tags on posts tagged "#emerson."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
ralph waldo emerson	201	201	0.20%
quote	170	614	100.00%
quotes	149	571	11.90%
life	59	211	10.21%
literature	37	103	2.07%
nature	35	45	0.48%
transcendentalism	33	33	0.02%

poetry	31	53	4.29%
ralph	23	23	0.01%
waldo	21	21	0.00%
inspiration	21	47	3.05%
philosophy	18	370	0.89%
truth	17	67	2.33%
wisdom	16	93	1.27%

Table 8.

*Tags on posts tagged "#heidegger."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
philosophy	56	370	0.89%
martin heidegger	29	29	0.01%
quote	21	614	100.00%
quotes	17	571	11.90%
dasein	14	15	0.00%
existentialism	13	92	0.10%
being	12	14	0.07%
being and time	12	12	0.00%
language	11	14	0.20%
nietzsche	8	587	0.16%
thought	8	19	0.43%
phenomenology	8	9	0.01%
poetry	7	53	4.29%
philosophie	7	13	0.02%

Table 9.

*Tags on posts tagged "#karl marx."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
marxism	27	31	0.02%
marx	24	26	0.02%
communism	20	24	0.03%

quotes	19	571	11.90%
capitalism	14	29	0.07%
quote	13	614	100.00%
socialism	12	18	0.02%
philosophy	8	370	0.89%
religion	8	23	0.46%
capital	7	7	0.01%
friedrich engels	6	6	0.00%
history	5	26	0.35%
das kapital	5	5	0.00%
politics	5	28	0.40%

Table 10.

*Tags on posts tagged "nietzsche."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
friedrich nietzsche	119	123	0.19%
quote	105	614	100.00%
philosophy	95	370	0.89%
quotes	91	571	11.90%
life	29	211	10.21%
filosofia	26	52	0.02%
love	24	96	15.87%
thus spoke zarathustra	21	21	0.02%
citazioni	15	35	0.33%
music	13	16	3.48%
frasi	12	18	0.30%
friedrich	12	12	0.01%
truth	12	67	2.33%
amor	11	16	0.75%

Table 11.

*Tags on posts tagged "#socrates."*

<b>Tag</b>	<b>Frequency (This Philosopher)</b>	<b>Frequency (All Philosophers in Corpus)</b>	<b>% Users in GQ Hashtag Corpus</b>
quotes	155	571	11.90%
quote	128	614	100.00%
philosophy	114	370	0.89%
wisdom	62	93	1.27%
life	54	211	10.21%
plato	52	59	0.11%
knowledge	31	47	0.36%
death	19	42	1.15%
love	18	96	15.87%
literature	16	103	2.07%
philosopher	14	25	0.04%
lit	13	70	1.93%
truth	12	67	2.33%
ignorance	12	15	0.12%

### **Multimodal Quotations**

I now turn to analyze those quotations that are "multimodal" in nature, containing both text and image. Such quotations are diverse, yet I draw attention to two different trends within them that, in their difference, give a sense of some of the further diversity of the ways that digital quotations and old media intersect.

**Quotation celebrates the book as a medium.** The prevalence of hashtags such as "#currently reading" point to the ways that quoting is tied to offline, print-based literacy practices. However, books are also present in quotation practice in an even more immediate form. Photo-type quotations tagged "#book quote" and "#book quotes" (BQ Photo Corpus) quotations often took the form of *photographs of books*.

In some cases, quotes consist of photographs of books marked with what might be called "private" markings, such as a brief (possibly unreadable) marginal note, shaky-handed highlighting or underlining done in pencil (Figure 3), or a simple five-pointed star. These are the sorts of notes that frequently populate the novels one finds in used bookstores. I refer to them as "private" because, in their simplicity and even hastiness, these signs suggest that their first purpose was to guide the attention of their creators during a private reading.



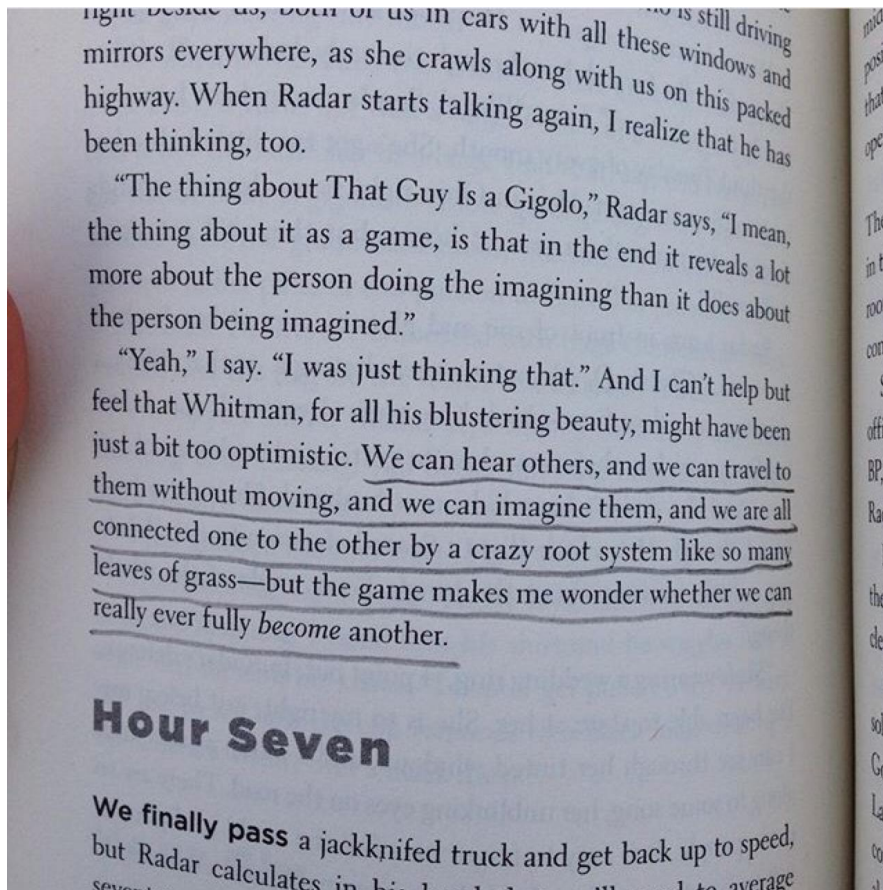


Figure 3. A "personal" marking made public online.

Photo-type quotes like this one are further proof that the practice of quoting on Tumblr and traditional practices of reading can be woven together in the lives of quoters/readers. Yet photos of pages can be far more than mere documents of reading practices. Sometimes they seem to anticipate (and in fact solicit) the attention of the wider Tumblr audience. A common tactic is not to mark the text at all but rather to *crop* an image of the text so it only contains the relevant words. In many cases, close cropping results in quotes that are long and narrow (Figure 4).

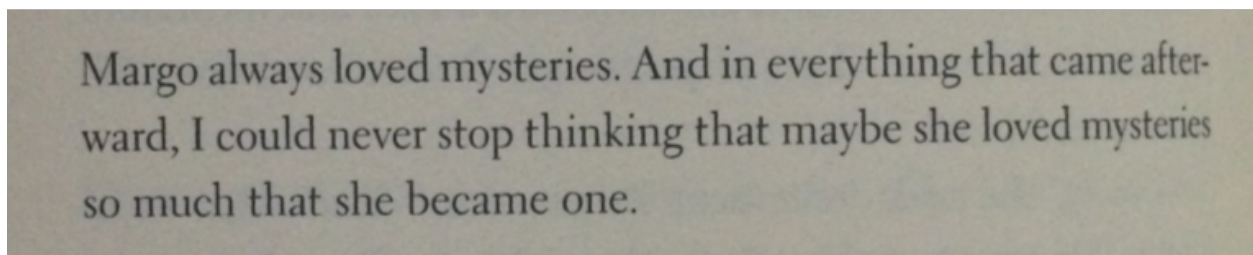
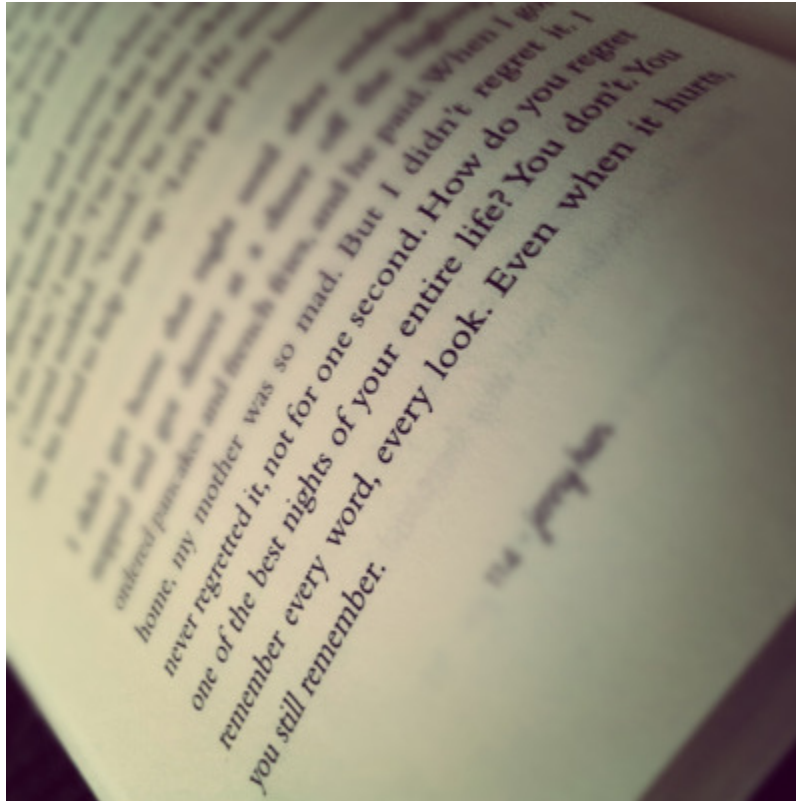


Figure 4. Cropping a text, likely via a photo-editing application.

In a sense, cropping and highlighting/underlining are equivalent, since they both draw attention to a particular passage. The former, however, is distinct in several respects. First, cropping aims to totally excise a quote from its surrounding context; this is unlike markings that make it possible to view as much as a page. If one of the features of a quote is that can be separated from its context, cropping makes this fact clear at the level of the materiality of the text. Second,

cropping is extremely precise, suggesting the use of either desktop or smartphone/tablet-based photo-manipulation software. Rarer than cropped photo-type quotes are quotes in which the non-relevant words of the text (those that are not part of the quote) have been blurred (Figure 5). As in cropping, the designer of the quote has made it impossible to read any of the surrounding text. And, as in the case of cropping, this is accomplished with computer-aided precision.



*Figure 5. Blurring a photo of a book.*

Cropping and blurring are paratextual markings that, unlike privately scribbled notes or underlinings, are first and foremost public; after all, these digital gestures leave no trace within the book itself. They are "functional" in that they reproduce a part of the book for other readers to consume on Tumblr, yet they may also aestheticize the book itself. This seems to be true in the example of the blurred quote, with its dramatic tilt. More rarely, photo-type quotes may also be even more obviously aestheticized. In one that I found, two pink flowers in full blossom have been carefully placed across the page, obscuring the unimportant words. In another, it is not flowers but dried leaves whose delicate withering echo the morbid theme of the quotation it adorns (Figure 6). (Notice how one leaf obscures the name of the speaking character, "the Korrigan.") These quotations oblige the goal of directing the reader's attention to a particular passage, but they are also decorative in nature.



"Fool is the man who denies happiness in favour of the past, for he will surely die wretched."

- *The Sons of Thestian*

# bookstagram # booklr # adultbooklr # b

290 notes

➤ 💬 ↺ ❤️

Figure 6. Aestheticized quoting: dried leaves.

Another quote, which has accrued over 7,000 notes, helps illuminate some of the underlying purposes manifested in such painstakingly artistic quotes. In this one (Figure 7), a book sits upon a stone at the edge of the water, no reader in sight.



“It is a condition of monsters that they do not perceive themselves as such. The dragon, you know, hunkered in the village devouring maidens, heard the townsfolk cry ‘Monster!’ and looked behind him.”  
—*Daughter of Smoke & Bone*, Laini Taylor

# it rained all day! # had to take this shot # c

7,188 notes



Figure 7. Aestheticized reading experience.

As this quote's tags suggest, this post is about "#reading" but it is also "#aesthetic," an

example of "#book photography." This blogger has even added certain descriptions of the photograph's style: "#pale," "#pale grunge," and "#foggy pale." (It is not tagged "#book quote," though it clearly features a quote from a book.) The quoted words themselves (a description of the condition of being a monster) do not in any obvious way echo the photograph. In this photograph, the book is more than a vehicle of information, and reading is more than the act of acquiring this information; instead, the book is a beautiful object worthy of photographic representation in its own right, and the act of reading (romantically, if a bit improbably, on a rocky shore) is a sumptuous experience.

Such photographs echo other trends in the ways that books and reading are represented on the site. Carefully posed still life photographs (Figure 8) represent reading through the *accouterments* that surround the practice, both necessary (glasses), comforting (a cozy blanket, a mug of tea), and decorative (a sprig of flowers, a bird's feather). Such decorations perhaps provide a "feminized" albeit rustic feel to the composition.



Figure 8. A book and tea.

Tumblr is awash in similar photographs. Sometimes they include a pack of expensive cigarettes (Dunhills, perhaps), but much more frequent is a cup of tea or coffee or a scented candle or even a sconce. Often their setting is in a bed, the perfect place to curl up with a good book, though quaint coffee shops are also popular settings. At certain points, "#book photography" and "#book quotes" do in fact overlap, as quotations too participate in this

aestheticized representation of reading. While it is known that Tumblr is a space for "fandoms," communities built around certain media *properties* (from *Harry Potter* to *Gilmore Girls*), aestheticized photographic representations of quotes, books, and reading itself are best understood of "meta-fandom," built around a media *form* rather than a particular property. To put it another way, these quotations are in fact celebrations of a certain kind of leisurely, print-based, and "deep" engagement with literature.

**Quotations are also transmedia artifacts.** Of the photo-type quotes bearing tags such as "#book quote," another one of the most common broad categories is those quotations that are illustrated with still images or short, looping videos sourced from film or television series. Often this quotation will "illustrate" a textual quotation, depicting the character who (in the novel, and in its adaptation) speaks the quoted words. On Tumblr, a "#book quote" and a "#movie quote" are not mutually exclusive. Clearly, Tumblr users appreciate certain works, from *The Hunger Games* to *Pride and Prejudice* not just as books but as "transmedia" narratives (Scolari & others, 2009), and the quote serves as a fannish object for celebrating various forms the story has taken, from book to feature film—or, in the case of works such as Austen's *Pride and Prejudice*—several feature films and television shows.

The creation of such quotations, however, must be contextualized within a broader cultural practice that is vibrant on Tumblr: fan editing, a process in which fans of a media franchise create different remixed versions of a media property. Such quotations sometimes bear the hashtag "#edit," evoking the fannish practice of remixing versions of official media properties.<sup>32</sup> To be clear, not all "#edits" are quotations. While some contain language culled from the source movie or show, others contain no words at all. The practice of fan-edits can be linked to specific fandoms. Fans of Jane Austen, for instance, use tags like "#austenedit" to categorize these fan-produced texts; tags like "#mine," "#my work," "#myedit," and "#creation" can be frequently seen on such fan edits, a sign of the ways that users take pride in themselves as creators of culture. Often fan-edits reference not just the character (such as the moody Mr. Darcy) but also the actor who played him (Matthew Macfayden, from the 2005 film adaptation, is a clear favorite). Some even note in hashtags the version of the film from which the stills or video has been sourced, such as "#p&p 2005." Such paratextual markings evince a librarian-like attention to the documenting of sources and important contextualizing information.

In some cases, fan-edits splice together written text and corresponding spoken dialogue: the words appearing beneath the GIF are those mouthed by the character or characters in the GIF, as if "closed-captioned." Others take more liberties in the way they combine film and text. In one, employing a collage technique, each frame features an interaction between Macfadyen's Darcy and Knightly's Emma. In one they dance, eyes locked. In the next, their eyes meet and look away. In one they laugh. In one she kisses his hand. In one he looks at her longingly. Only the middle frame is emblazoned with the words of Darcy, "You have bewitched me, body and soul." This last line, however, is not actually from Austen's novel but from that particular film adaptation.

Often it is not obvious whether a particular quoted dialogue is drawn from the source text or one of its many re-imaginings. Yet sometimes it is, as when a short looping GIF accompanies a paragraph of what is clearly prose. One quote tagged "#p&p 1995" consists of four GIFs drawn from the 1995 BBC version starring Jennifer Ehle and Colin Firth. Read left to right and top to

---

<sup>32</sup>See, for instance, Phillips (2011).

bottom, these GIFs capture only a few seconds of the film. As the first one begins its loop, Lizzy Bennet's smile cracks; from there, Darcy's eyes turn downcast and he walks out of the room, but not before casting one more look upon her. In the final GIF, Lizzy's smile is gone. This collage of wordless expressions is explained by the accompanying text:

*As he quitted the room, Elizabeth felt how improbable it was that they should ever see each other again on such terms of cordiality as had marked their several meetings in Derbyshire; and as she threw a retrospective glance over the whole of their acquaintance, so full of contradictions and varieties, sighed at the perverseness of those feelings which would now have promoted its continuance, and would formerly have rejoiced in its termination.*

In bold, next to this text, are the words "Chapter 46." This quote—and it is tagged "#austen quotes"—knowingly combines both novelistic prose and filmic image. Considered as a kind of quote, this creation bespeaks a deep familiarity with *Pride and Prejudice* the book and with *Pride and Prejudice* the movie, using parts of each as necessary. Quotations like this one represent a logic that is contradictory to those quotes featuring aestheticized depictions of a physical book, valuing not a specific media form but a narrative that exceeds any single one of its media-specific manifestations. Yet this quotation is also evidence of the ways that "deep" attention may take on new forms in a digital, transmedia context; arguably, the ability to weave together media represents an engagement with the source texts that is *deeper* than merely making one's way through the novel.

### **Other Signs of Practice**

To conclude the analysis of this chapter, I explore two aspects of quotations as a "textual ecosystem." First I explore the prevalence of misquotation before turning to the exploration of some of the ways that quotations can occur within (and circulate between) Tumblr blogs. Understanding the phenomenon of misquotation is important because it represents a way of quoting that is obviously not "deeply" attentive to the source. (This is not to say that a quoter might not *understand* or *feel* the quote in ways that could be considered "deep.") The ways that quotations circulate shed light on the ways that bookish quotations on Tumblr may receive different kinds of attention from different types of users.

#### **Misquotation.**

*Misquotations do occur on Tumblr.* Out of 8 fabricated or misattributed quotes I investigated from Wikiquotes, I was easily able to find examples of six of them on Tumblr. Some of these misquotations seem rather harmless. On Tumblr, the likes of Voltaire, Emma Goldman, and Marcus Aurelius are said to have spoken or written words which they in fact did not. Perhaps most disturbing is the fact that a white supremacist named Kevin Storm can be found ascribed to Voltaire; one example I found received over 30 likes or reblogs. In fact, misquotations can become extremely popular.

Analysis of the Philosopher Corpus revealed that for certain philosophical authors at least one of their top 5 most popular quotations were misquotations of one form or another. Camus did not write this:

*In the midst of hate, I found there was, within me, an invincible love. In the midst of tears, I found there was, within me, an invincible smile. In the midst of chaos, I found there was, within me, an invincible calm. I realized, through it all, that... In the middle of winter, I*

*finally found that within me there lies an invincible summer.*  
though he did write the final line, which appears in his essay "Return to Tipasa" (Camus, 1967, pp. 169<sup>33</sup>; cf. Gaetani, 2015). This misquotation received over 2,400 likes or reblogs. Oddly, the blogger who created it attributed it not just to Camus but to a specific work: *The Stranger*. The misquotation, however, betrays a less than "deep" attention to this particular philosophic novel.

Other misquotes are even more ambiguous. Any reader of Marx, the revolutionary economist, would likely have a hard time believing that he wrote a quote that begins: "Surround yourself with people who make you happy." In fact, something like this was perhaps written by a different Karl Marx, a composer.<sup>34</sup> However, the version of this quote that received over 2,000 likes or reblogs did not make this clear. Another blogger also posted this quote, but added as a comment, "Karl Marx said this?"

Sometimes "misquotations" are not totally false but rather have simply been excerpted cleverly but disingenuously from the source text. I was suspicious of this popular quote, again attributed to Karl Marx:

*patience – the first condition of learning anything*  
Moderately popular, it received over 300 likes or reblogs. Karl Marx did indeed write these words in a letter to Friedrich Sorge—yet the purpose of this letter was not to bestow upon his intellectual comrade beatific wisdom concerning the virtue of patience. Rather the purpose of the letter, composed in 1881, was to complain of the way their work *Das Capital* had recently been discussed in British periodicals and to deride a recent work by British writer Henry Hyndman. In this letter, Marx is clearly aggrieved that Hyndman's work is so indebted to *Das Capital* but does not mention Marx by name. Worse, Hyndman's book is simply not very good, and so it is a copy which insults that which it emulates:

*...[H]is little book – so far as it pilfers the Capital – makes good propaganda, although the man is a "weak" vessel, and very far from having even the **patience – the first condition of learning anything** – of studying a matter thoroughly.*<sup>35</sup>

Marx's claim about patience is bracketed off, an observation rattled off in the process of levying a potent insult against what he regards with sneering condescension as an "amiable middle-class writer." The original passage is not really about "patience," as the Tumblr quote makes it seem, and in order to achieve this effect the creator of this quotation has done a certain violence to the original language of the passage. Not only has a noun phrase within a subordinate clause (beginning "although the man is weak...") been posited as a complete (albeit verbless) sentence, the word "patience" has been separated from the prepositional phrase with which it forms a larger noun phrase ("patience...of studying a matter thoroughly"). This effects a tonal change too, as the original passage's vitriol is erased. This transformed quote gives the impression that Marx *qua philosopher* wrote about "patience"—not technically incorrect, but certainly incomplete. A specific argument becomes slyly recast as an inspirational maxim.

Other quotations are much less ambiguous. They are simply incorrect. One of the most popular quotes attributed to Socrates (receiving nearly 3,500 likes or reblogs) is this statement

---

<sup>33</sup> In E.C. Kennedy's translation: "...there lay an invincible summer...".

<sup>34</sup> According to wikiquote: [https://en.wikiquote.org/wiki/Karl\\_Marx](https://en.wikiquote.org/wiki/Karl_Marx)

<sup>35</sup> In my Google searches, one of the topmost hits was this letter transcribed on the Marxist Internet Archive ([https://www.marxists.org/archive/marx/works/1881/letters/81\\_12\\_15.htm](https://www.marxists.org/archive/marx/works/1881/letters/81_12_15.htm)). I suspect the quote may have been extracted from this site, as other versions of this passage I found via Google Books use an en rather than em dash (cf. Marx & Engels, 1942).



about relationships:

*Some people put walls up, not to keep people out, but to see who cares enough to knock them down.*

To the best of my knowledge, the words of Socrates, as they are related in the works of Plato, express no such sentiments. Neither does Plato record him as saying this:

*Life contains two tragedies. One is not to get your heart's desire; the other is to get it.*

This quote received even more likes and reblogs, nearly 7,000.

Misquotation occurs on Tumblr, and while it is important not to assume that the misquotations quotations observed above are *typical* of quotations, it is clear that at least in some cases quoting does seem at odds with the "deep" attention of traditional reading. "Reading" Socrates via quotations on Tumblr can be utterly disconnected from reading his words in dialogues such as *Meno* or *Gorgias*.

This is not to say that the Tumblr users who share and reblog these misquotes do not understand or feel them deeply; it is important not to speculate about the invisible cognitive or affective encounters readers have with these texts. Insofar as quotations may provide solace, guidance, or inspiration, perhaps whether they are "real" is beside the point, as is whether or not a quotation emerged from a "deep" reading of a longform text.

*Vernacular bibliography: misquotation's antidote.* On the other hand, it is possible to also discover on Tumblr what might be thought of as the *opposite* of misquotation—quotation practices in which quoters carefully document the link between the quotation and the source text, often a book.

Sometimes individual Tumblr users seem to actively wage war against a specific misquotation. Take, for instance, the quotation by Evelyn Beatrice Hall, wrongly attributed to Voltaire. Certain versions of this quotation on Tumblr do not simply correct the misattribution; rather they also work to clear up the confusion by noting the original book from which this quotation originates. It is "[f]rom her biography of François-Marie Arouet (Voltaire)," according to one blogger. According to another, it is "a summation of Voltaire's beliefs on freedom of thought and expression." Yet another notes that Hall (writing under the pseudonym S.G. Tallentyre) notes the date of the original work in which these words appear, 1,907, and explains that Hall invented something as "an example of something Voltaire might have said." Not only individual bloggers but also institutions have chimed in; the official Tumblr blog of Oxford University Press devotes a page to documenting the correct source, in the process quoting an extended passage from Hall's book *The Friends of Voltaire* from which the quote originates. Yet one blogger, in repeating the quote, has gone a step further, linking to an electronic version of *Friends of Voltaire*.

In the case of this frequently mis-cited quote, users engage in what could be called "vernacular bibliography"—referencing and even citing textual sources within social media posts in order to explicitly resist disinformation regarding the words of a particular author drawn from a particular book. In fact, referencing sources is an important part of quotation practice on Tumblr, especially the quotation of books. In a sample of 100 randomly chosen source fields from the BQ Source Corpus (representing 100 different users), 73 of them mentioned both an author and a book (e.g. "Delirium by Lauren Oliver"). Several also documented the character speaking the quotation in the case of reported speech (e.g. "Toru Watanabe from Norwegian Wood by Haruki Murakami"), and one noted the chapter and page number. Referencing an author does not guarantee veracity, as the quotation wrongly attributed to Camus' *Stranger*

makes plain. Still, when citing books on Tumblr (at least with the tag "#book quote" and "#book quotes"), Tumblr users generally provide at least minimal bibliographic information.

### **Quotations Within Blogs.**

*"Bookish" Blogs.* In the case of blogs that post book quotes, some blogs can at times be considered "bookish" in that they are primarily about books (if not exclusively quotes from them).

The bookish nature of blogs can also be seen in the ways that Tumblr users name their blogs and their blog descriptions. Out of the 7,261 blogs in the Usernames Corpus, more than 9% of them had blog names referring explicitly to books or reading—names like "book-under-the-clouds" or "bibliofantasia." Out of a sample of 1,000 blogs in the Blog Description Corpus, 23% of their descriptions mentioned books or reading. Sometimes these blog descriptions make it clear that books are the only (or at least main) topic of a blog:

...quotes from all the books in my life.

In a bibliophile and that's about it

In other cases, they reference a variety of media or even a more diverse set of interests that may appear on the blog:

I love anime, books, pokemon, and netflix

Industrial Design, Sewing, Books, all Music, Tv Shows, Films, Food, Pusheen<sup>36</sup>, Life, people, travel, culinary, design, books, etc.

In such cases, bookishness is not presented as incompatible with other non-literary interests or content.

*Bookish blogs vs. non-bookish audiences.* Those who consume book quotations are not as "bookish" as those who blog them originally.

Out of a sample of 826 users who had consumed a quote by liking or reblogging it (Audience Usernames Corpus), only 2% of these users' usernames contained a reference to books or reading, compared to 9% for those who *posted* these same quotes; this difference was statistically significant ( $\chi^2=43.6$ ,  $p<.001$ ). I interpret this as a sign that those who post quotes from books are more "bookish" than those who consume them. By posting blogs from books, book quoters are introducing quotations from books to Tumblr users who are not as "bookish" as they are. Considered as a group, those who post quotes from books may be thought of as "sponsors of literacy" (Brandt & Clinton, 2002, pp. 349) who "underwrite" less bookish users' access to the language of books.

To be clear, what is true of the consumers and producers of book quotes may be true of other media (e.g. fashion photography, fan fiction). These discrepancies between those who post book blogs and those who consume them gets at a fundamental aspect of Tumblr as a textual ecosystem: because of reblogging, texts that originally appeared in a certain kind of blog may end up in blogs that are very different in content and style. And the varied contexts in which a single post appears can hint at very different ways of "paying attention" to a book quote. Consider the following "focal" quotation:

*Longing is the agony of the nearness of the distant.*

The quotation is attributed to Martin Heidegger; his full name is given, as well as the quoted

---

<sup>36</sup>A cartoon cat.

work: *Who is Nietzsche's Zarathustra?* I traced this quotation to its original blog, which seemed noteworthy in several ways. First, while other blogs that I had encountered were made of a (sometimes riotous) collage of words, images, GIFs, and other media, this blog is mostly filled mostly with quote-type posts. While blogs sometimes have colorful or even animated backgrounds and designs, this one is tidy, minimalist, even austere: black letters on a white background. What stood out most, however, was the rarified collection of authors it quotes, many of which are likely more obscure (or, in any case, "academic") than Heidegger himself: Helene Cixous, Gerard Bruns, Augusto Boal, Nathaniel Mackey, Jed Rasula, and Lorine Niedeker are good examples of the sorts of sources this blog draws from. The fact that the quote from Heidegger links to the full document on the archive JSTOR serves as further evidence of an academic disposition, and indeed an "about" section makes plain the academic credentials of this blog's author. Quotations from academic authors are not the *only* content of the blog, but they are certainly the bulk of it. The sorts of authors so popular in the practice of book quoting—John Green, Rainbow Rowell—were conspicuously missing, and in their place extended passages from Adorno. In short, this blog stood out to me as perhaps the most overtly academic quote blog I had come upon.

Many of the quotes from this blog have achieved a modest popularity (fewer than 100 "notes"). However, this quote from Heidegger somehow achieved an order of magnitude greater virality, collecting over 1,000 "notes." While not all of these are reblogs, many of them are; these Heideggerian words, reblogged, now reside in many contexts other than this academic quote blog. The quote, in other words, has gone on a journey, a journey that has taken it to places that are very different than its original home. Below I present a "context collage," consisting of a brief description of the most recent post from 25 different tumblr blogs that reblogged this quote. As described in Chapter 2, the goal of this collage is to present a sampling of the different contexts in which this quotation has appeared.

A black-and-white picture of a girl on a bicycle, looking up the steeply-climbing San Francisco street. An authorless quote about longing. A dense paragraph attributed to Guy Debord. A joke about trying to be more like Beyoncé. A quotation from a poem by Charles Bukowski. A photo-montage from Kubrick's *Eyes Wide Shut*; Nicole Kidman extricates herself from the confines of her dress. A quote by George Bataille. The cover of an obscure self-released Polish electronic music album. A black and white photo of hands. A photograph of people's heads, seen from above. A photograph of salsas: orange, red, and tomatillo-green, in wooden bowls on a colorful tablecloth. An off-color joke. A photo-montage of a fair; in the first, the cow seems to be smiling. A photograph of an elephant. A movie gif: Kiera Knightly in period dress somberly twists in a swing. A quote from *The Vow*, a romantic drama starring Channing Tatum and Rachel McAdams. A Soundcloud link to a track by an Australian producer. A quote from Jude the Obscure. A photograph of two young men clad in black; one of them has a tattoo and a wine bottle. A quote from *The Glass Menagerie*. A YouTube link to a song by an Eastern European punk rock group. A close-up of dulce de leche chocolate cupcakes. A (fashion?) photograph of a woman, shadows crossing her face. A photograph of a cassette player. A photograph of a woman; a butterfly has landed on her wrist. A photo-montage of horror-movie creatures resting on set.

This is only a small fraction of this quotation's "neighbors" on various blogs around Tumblr. The most important lesson to be learned from this motley array is that quotations end up in radically different contexts, some of them more surprising than others. Certain neighbors seem almost

totally unrelated. Heidegger's meditation on longing has little to do with cupcakes, or a fair, or an elephant. A quote from *The Vow*, a major-market romance based on a bestselling novel, is very different than the obscure and academic sources favored by the originally-posting blog. Likewise, the originally-posting blog does not tend to feature authorless quotes; each quote's author, from Robert Duncan to Walter Benjamin, is meticulously documented.

On the other hand, other neighbors seem more "related," so to speak, and not just because they are also quotations. One might imagine that a reader of Heidegger is also a reader of Debord and Bataille. Searching Google Scholar for "Heidegger"+"Debord" returns a long list of academic works that discuss these thinkers in tandem. "Heidegger"+"Bataille" is a less frequent combination, yet certainly all three share some obvious similarities. They are white, 20th century, male, European intellectuals who studied literature and philosophy.

A quoter who quotes Bataille and Heidegger says to the reader "I am interested in European philosophy." And in doing so it says something more basic: "I know something about European philosophy, at least enough to know that both Bataille and Heidegger are both European philosophers." Blogging a quote from Heidegger and a picture of a dulce de leche, however, is a fundamentally different semiotic gesture altogether. While it would be wrong to speculate too much about the degree to which the user who reblogs both of these really "gets" Heidegger, it is fair to say that a blog that contains both Heidegger and food photography is more diverse than the one that contains almost exclusively philosophical or literary quotations.

While it is possible to create a quote-filled blog on Tumblr that is overtly academic, a quotation that itself seems academic in this context may seem less so as bounces around the site. This relates to the matter of attention. The blog of the user who originally posted this quote is itself deeply attentive not just to books but to a certain kind of book. But this post becomes reappropriated, landing next to pictures from movies and food and a variety of other media; many if not most of these other contexts suggest that the reblogging users pay attention to much more diverse types of media.

Once again, it is important not to speculate too directly about the character of particular quoter's understanding of quotes. A Tumblr blogger who reblogs a few choice words from Heidegger amidst a kaleidoscopic array of humorous memes and food photography may understand this fragmentary philosophical text just as deeply as the more obviously academic blogger. What is visible in these blogs is not attention, exactly, but the *performance* of attention. One can present oneself as "deeply" engaged with a certain bookish domain at the expense of other domains, or one can sample promiscuously from books and other domains: *I pay attention to everything*, this blog implies, *and I draw no firm distinctions*. Such a "promiscuous" blog is reminiscent of Bakhtin's (1984) concept of the *carnival*: "all things that were once self-enclosed, disunified, distanced from one another" meet and commingle during this special time during which hierarchies vanish, "the sacred with the profane, the lofty with the low" (pp. 251). By this light, a user whose blog pays sole attention to books (or perhaps a more specific category: philosophical works or *avant-garde* literature) tries to maintain the "self-enclosed" (and certainly "lofty," even "sacred") status of literature and specific genres within it. Yet Tumblr's carnival prevails as literature finds itself wedged between some very unlikely bedfellows (or blogfellows). Another word for "hyper" attention, then, might be "carnavalesque" attention.

### Discussion

What does it mean to read something "deeply"? This study began with a rather stark division that assumed both that old ways of reading (e.g. "getting lost in a novel") are deep while

new, digital forms of reading tend toward the opposite pole, demanding "hyper" attention. While there is perhaps some truth in this binary, this chapter has revealed quotation as it occurs on Tumblr to be anything but a single, unitary practice. Quotations are diverse, and their diversity suggests ways of quoting that seem "deep" and others that seem less so.

To be fair, the critic who already looks askance at newfangled forms of reading by now will have found more evidence to confirm a suspicion that whatever new ways of experiencing literature, philosophy, and other texts has emerged on Tumblr, they are simply no replacement for "real" reading. Misquotation especially is not just an occasional infelicity but a fundamental aspect of quotation on Tumblr, and so the critic has good reason to doubt that the experiences of quoting and consuming quotes assigned, sometimes dubiously, to certain writers and thinkers is as "deep" a literacy experience as simply reading them. On the other hand, some quoters rail against misquotes. Others translate into hashtags the words one learns in college or graduate school. Others dutifully post the links to the full texts, connecting Tumblr to more traditional archives of knowledge. Quotes posted by a blog that posts almost exclusively from books or academic texts—arguably another sign of "deep" attention to such texts—may be reblogged by other bloggers who themselves do not profess or display such a bookish disposition, or who post miscellaneous ("carnavalesque") content. And quoting at least *can* be connected to more traditionally valued regimes of reading; it is not necessarily something done *in lieu* of reading but can occur in the process of reading itself.

However, there is one way, undiscussed so far, in which quotation as a literacy practice differs from certain kinds of "deep" reading, especially the "critical" kinds that go on in schools and universities. Earlier in this chapter I briefly mentioned that in hashtags book quoters will sometimes document their reactions to quotations. To be more specific, these reactions are almost always *positive*. The BQ Hashtag Corpus abounds with examples in which the quoter's affection for the quote, or the book from which the quote was sourced, is obvious. For example:

i love this

i love this quote

love this quote too much

These examples were only a few of the over 50 tags found in this corpus by searching for those containing the string "love this". By contrast, only one contained the string "hate this" — "I hate this." The point is obvious, yet it needs to be made explicitly: quoting a text is generally a way of affirming a quotation. While observing the hashtags that users tend to use to describe quotations, it is worth keeping in mind the kinds of terms that they do not use. I came upon no instances of quotations attributed to Heidegger, for instance, that mentioned the connections between his philosophy and Nazism. On a more general level, quoters tend to share quotes that are examples of "#truth," not of "#falsehood." Thus, while this chapter has suggested that there are ways in which ostensibly academic texts as well as academic discourses manifest themselves in quotation practices on Tumblr, these practices may be best characterized not by a lack of depth but by a general positivity toward the quoted texts. Understanding the reasons for this positivity, however, will require leaving the textual ecosystem of Tumblr and going inside the perspectives of quoters themselves, the work of the next chapter.

## CHAPTER 4

### QUOTERS IN THEIR OWN WORDS

While Phase I concerned *digital texts*, Phase II supplements this view with a properly ethnographic investigation of *digital quoters*, exploring the role that quotation plays in the lives of those who create and share them. In particular, this phase looks at those quoters who have used the "#book quote" or "#book quotes" tag, with goal of answering the second main research question about how quoters understand the value of quotes and the purpose of quotations that involve old, print-based media. It further explores how quoters understand quotation as fitting into their larger experiences of literacy.

Unlike Phase I, which aims to present quotation culture on Tumblr in its complexity (observing some of the different and even contradictory tendencies manifested within it), Phase II was an attempt to understand the practices of those who *do* seem to quote "bookishly," with the caveat that they likely do not represent quoters in general. The central goal of this Phase is to understand in some detail the ways in which quoting from texts may be symbiotically intertwined with "deep" reading. It pursues this goal by also trying to account for quoters' own understanding of quotation's value in their lives and how different forms of attention may be necessitated by different purposes for quoting and different uses of quotations. Quoters use quoting as a way of paying attention to books in ways that conform to and also complicate what the term "deep" attention typically means. Furthermore, I explore the ways that quotation is a key part of practices of self-expression and self-care, fundamental purposes of literacy that represent another non-cognitive way of experiencing texts "deeply."

### Data and Methods

#### Data

The main data sources are the results of a questionnaire and follow-up, on-line interviews with quoters.

**Questionnaire: Reaching Out to Tumblr Users.** To find out quoters' perspectives on quotes and quotation (as well as certain demographic information to contextualize their responses) I began with a questionnaire (Appendix A). This instrument was designed to ascertain (1) respondents' age and gender, (2) the degree to which their consumption of quotations supports their consumption of print texts (specifically, whether they use quotations as a way of discovering books to read), (3) their general thoughts on the purposes of quotation, and (4) the qualities that they think make a quotation quotable. The question about the relationship between quoting and the discovery of print text (2 above), most directly gets at the question of attention. With the other questions I hoped to gather information about who quoters are, their reasons for quoting, and their understanding of quotations. While these matters are not directly related to attention, they provide important contextual information. I assumed that any discussion of how quoters pay attention to books via quoting would need to be contextualized in a general understanding of quotation as a practice.

**Participants.** Prospective recipients of the questionnaire were those users whose blog

names appeared in the Unfiltered BQ Corpus; in other words, they were bloggers who had used the tag "#book quote" or "#book quotes." One challenge of online research is that not all participants are equally present on the site, and may have stopped using it long ago. Pilot interviews suggested that simply sampling randomly from a list of quoters would lead to an extremely low response rate. To deal with this problem, I re-organized quoters in terms of their most recent post on Tumblr (also ascertainable via the Tumblr API) and sent the questionnaire to them (a principle of "last-posted, first-surveyed").

I sent the questionnaire, along with a brief introduction and consent documentation, as a link embedded in a Tumblr "chat" message. A small minority (<5%) of users had set their accounts to block attempts at chatting; I was not able to send the questionnaire to them. In the end, I sent 538 links to the questionnaire and received 157 responses, a response-rate of 31%. Looking through the responses, I noticed that, in answering to a question that asked respondents to describe how they go about finding quotations to post on their blog, not all users mentioned quoting from books. I decided to categorize the 88 users who *did* mention quoting from books as a way of getting quotes for their blogs as "book quoters" (or "BQs"). I focus on book quoters, but did not throw out the other respondents (whom I call "other Tumblrs" or "OTs"), whose answers often suggested familiarity not just with Tumblr in general but also with quotation.

Due to my research protocol, I was not able to send the questionnaire minors, and so my sample was likely biased against younger users.

**Follow-up interviews.** I conducted follow-up ethnographic interviews (Spradley, 2016) asynchronously through the Tumblr "chat" feature. Questions began by expanding directly on respondents' questionnaire answers which often led to more general discussions of quotation as a practice. These interviews were crucial for getting at the highly-idiosyncratic ways that individuals practice quoting. I was able to conduct follow-up interviews with 27 out of 66 respondents who signaled their willingness to be contacted again. (In addition, 10 of those 66 were contacted but did not respond to my request.) Some interactions were short, consisting of a single response to an often very specific question about an answer from the questionnaire, often merely to clarify a questionnaire response. Others consisted of several conversational turns, while several became more free-flowing conversations taking place over the course of several days or even weeks.

One limitation of the data analyzed in Phase I is that it only illuminates quotation as an "online" practice. But any such practice is also at some point embodied, carried out by real people in the physical world, interacting with some kind of hardware. Of course, I could not see this happen—yet such blindness is a limitation of ethnography of reading in general, insofar as much reading is by its nature private. My solution was to ask certain questions that at times seemed to strike subjects as rather banal: do you enter quotes on your phone or your computer? When you are reading, do you have your device nearby so you can post quotes as you read? Such questions, essentially "mini-tour" questions (Spradley, 2016), get at the embodied practice of quoting as it relates directly to different ways of paying attention to books via quoting.

Through what Spradley would call "native language questions," I was able to solicit respondents' own vocabulary for describing quotes and quotation. I also used these follow-up questions as a way to validate ethnographic intuitions as well as to complicate my understanding and generate further questions. For instance, I would paraphrase my own preliminary analyses of the purposes of quotation in general or specific types of quotations, asking the respondents whether they agreed with my tentative findings derived through conversation with other quoters

as well as other analysis of data (what Spradley calls "verification questions"). My interlocutors were not shy to disagree with me, posing alternative explanations for specific findings.

I found that this process was especially helpful for surfacing bloggers' own rich, complex theories about quotation and its purposes—a topic that some were obviously excited to have been asked about.

**Blog Description Corpus.** I once again used the Blog Description Corpus, this time for a simpler purpose: Tumblr users sometimes list their ages as well as their names, which can serve as a proxy for gender. (While many Tumblr users veil themselves behind anonymizing usernames, this is not always the case.) This corpus thus can be mined for demographic information about quoters.

## Analysis

**Analyzing questionnaire responses.** Questions were of different types and required different techniques of analysis.

- In the case of the open-ended question asking respondents to name the "most common ways that you go about finding quotes to post," (see Appendix A, question 2), I first noted answers that made reference to finding quotes in reading or in books. From this group I made a broad distinction between those respondents who primarily draw quotations from books and those who do not (what I call "reader-quoters" versus simply "bloggers"). I then looked for patterns in the sources other than books mentioned by "bloggers," as well as whether or not reader-quoters described getting quotes from sources other than books in addition to books.
- In the case of open-ended questions for which I had no *a priori* categories (for instance, the question about the characteristics of a good quotation, [Appendix A, question 4]), analysis took the form of cycles of qualitative interpretive coding. Saldaña (2015) describes two main cycles of coding. In the first cycle, the researcher identifies and names patterns, in the process dividing the data into key analytic units. During this cycle, I often began by noticing concrete repetitions in the language of respondents before generating slightly more abstract descriptive codes. For instance, after noticing that respondents seemed to often say that quotes must "resonate" with them, I took this vocabulary to develop the code RESONATE; I also used this to label other statements that I considered to be essentially similar, such as that a quote must be "relatable to my own life." In the second cycle, I synthesized these low-level patterns into higher-level, interconnected themes, producing a taxonomy that begins to structure an overall interpretation of the data. For instance, after noticing that quoters suggested that quotations need to be linguistically interesting (code: POWERFUL LANGUAGE) as well as short (code: BREVITY), I grouped these together as formal characteristics of quotes (code: FORMAL), which in turn I juxtaposed to another nested branch of codes referring to the qualities of quotations that were related not to formal but to semantic aspects (code: MEANING). Each open-ended question necessitated the development of codes (see Appendix B). As Sandaña (2015, pp. 4) notes, coding is not an end in itself but rather a "transitional process between data collection and more extensive analysis." In describing the data in this chapter, this "more extensive analysis" included at times analysis of which codes I used most frequently, signaling the major trends in the data. But, more than noticing trends, I have reproduced significant quantities of quoters' own



words (hence this chapter's title), an approach that allows me to further witness and analyze the complexities that cannot be fully captured in a mere code.

- Answers to the Likert-scale question about how frequently respondents discovered books via quotations were tabulated so that I could see most common answers as well as the difference between the number of respondents who signaled that they never do this and those who signaled that they do, even if rarely.
- Answers to a question requiring a numerical answer (age of respondents) were averaged; the mode was also calculated. A question on gender was open ended (allowing non-binary answers), though almost all answers fell into a male/female gender binary, and so a male/female ratio was calculated.

**Analyzing follow-up interviews.** Once again, qualitative coding was used to surface main themes in responses. Most of the main codes developed to describe the questionnaire data remained relevant, though I did need to develop further "process" codes to categorize quoters' detailed description of their quoting processes.

**Analyzing Blog Description Corpus.** It was important to link quoters' purposes for quoting to basic demographic information about the quoters. My analysis of the Blog Description Corpus was limited to the further exploration of demographic information (age and gender). Using a regular expression (a computational tool for pattern matching), I extracted strings of two sequential numbers representing age (e.g. "17," "34"). I manually removed false positives (numbers that clearly, based on context, did not refer to an age) and calculated the mean of these average ages.

Users also often list their names; this can serve as a proxy for gender. I extracted possible names using regular expressions. One captured the first word of the description, since users often include their names before any other details. Others looked for patterns like "My name is Kyle." Using the Python NLTK Names Corpus, which contains a large number of typically male and female names (with some overlap between the two lists), I checked each possible name, noting the gender of those names that were, according to this dictionary, unequivocally male or female.

**Exploring "nonhuman" Tumblr users.** Midway through this phase of research, I noticed that institutions, not just individuals, post quotations on Tumblr. These include literary agencies and publishers. To account for this violation of my own assumptions of who is on Tumblr, I investigated the Tumblr blogs of the so-called "Big Five" publishing houses: HarperCollins, Macmillan, Penguin Random House, and Simon & Schuster, collecting and analyzing samples these blogs to see if publishing houses posted quotes.

## Findings

### Who Quotes?

The respondents to my original questionnaire, those who use the tag "#book quote" or "#book quotes," are in some ways not a homogenous group.

Most of them (82 out of 126, or 65%, of those who answered the relevant questionnaire question) identified themselves as what I call "reader-quoters"; they described directly excerpting quotes from favorite books and books that they happened to be reading. Most other quoters, however, described sourcing quotes from a variety of "secondary sources," including reblogging

from one of the "literary-specific blogs" they follow on Tumblr (Blogger42, f, 28, Ques)<sup>37</sup>, sourcing quotes from Goodreads (Blogger49, n, 21, Ques), or searching for them using "reliable google" (Blogger76, f, 25, Ques), a turn of phrase that suggests the search engine's ubiquity.

Yet the boundary between quoting-reading and other quotational practices is not a firm one. Of the 157 respondents I labeled "reader-quoters," 35 of them (42%) described finding quotes from books but also from at least one other source. RQ47 (f, 20, Ques), for instance, reported that she finds quotes through reading books *and* directly from other blogs and even Goodreads. Another reader-quoter (RQ39, f, 31, Ques) cites snippets from her own reading but also finds quotes, she said, by "searching authors that I read before to fin[d] quotes from other books they wrote." Furthermore, books are not the only media from which reader-quoters extract quotes. RQ6 (f, 20, Ques) described gathering quotes from books as well as music and movies. RQ40 (f, 18, Ques) finds her quotes from "read[ing] widely and watch[ing] lots of good movies and drama." For Tumblr users like these, books are just one among various media forms to be quoted.

**Gender and age.** Despite the diversity of practices that may be involved in quoting from books, those who post "#book quotes" are homogenous in terms of gender. Respondents to the questionnaire overwhelmingly identified as female. Of the 99 respondents who answered a question on gender identity, 91 described themselves as female, 5 as male, and 3 outside of a gender binary. The same was true when looking specifically at the 74 reader-quoters who answered this question: 69 of whom identified a female, 3 as male, and 2 outside of a gender binary. Analysis of the Blog Description Corpus, descriptions of blogs where the "#book quote" or "#book quotes" tag had appeared, confirmed this gender imbalance; 101 of the names listed in the sample of 1,000 blog descriptions were unambiguously female, while only 17 were unambiguously male. That an online culture devoted to books exhibits a level of gender disparity is perhaps not surprising; after all, women and girls in general are more likely to be readers (Perrin, 2016). However, the exact level of gender disparity here is striking, especially in light of the fact that Tumblr in general is more or less gender-balanced ("Demographics of social media users," 2015).

Like Tumblr users in general, however, questionnaire respondents were young. Their average age was 24.21 (n=104), and the average age of reader-quoters was 23.7 (n=77). Since the questionnaire necessarily excluded those who were not at least 18, it seems reasonable to assume that this number overestimates the age of quoters on Tumblr. Indeed, analysis of the Blog Description Corpus revealed an average age of 20.5, with the three most popular ages listed being 19, 18, and 21.

This preliminary finding frames the more detailed, ethnographic ones that followed. While this chapter explores what quoters value in quotations, the purposes that motivate them to quote, and the ways that quotation connects to their larger experiences of literacy and forms of literate attention, it also considers the degree to which quotation is at least on Tumblr a gendered practice.

Though my research methodology did not explore Tumblr users' nationalities, it became apparent to me during follow-up conversations that participants were a surprisingly international

---

<sup>37</sup>Each respondent is referred to by a unique id, a gender ("f" for female, "m" for male, "n" for non-binary), an age, and, usually, a source ("Ques" for questionnaire, or "Chat" for follow-up chat). The unique names of reader-quoters begin "RQ"; others begin "Blogger."

group. I came to learn that those I interviewed came not just from the United States but also Finland, Sweden, the Philippines, Switzerland, and Brazil. These revelations hint at the ways that book quoting—even in English—may in fact be a global practice.<sup>38</sup>

### Theorizing Quotability

In analyzing the questionnaire questions about the qualities of a good quotation (Appendix A, Question 4), two main patterns emerged: a good quotation is one that is (a) linguistically virtuosic and/or (b) both personally resonant and universally applicable. These facts suggest some of the ways that quoters "pay attention" to quotations.

**Powerful language.** 33 reader-quoters, in addition to 5 bloggers, referred directly to the ways that quotations contain language that is somehow more artful than everyday language or average prose. They contain "witty words or puns" (RQ23, f, 25, Ques) or "a play on words" (RQ44, f, 19, Ques), "poetic wording" (RQ44, f, 19, Ques), "proper syntax" (Blogger54, f, 21, Ques), "excellent punctuation" and "variation of phrase" (RQ69, f, 21, Ques), "vivid imagery" (Blogger24, f, 18, Ques) that "goes above the typical everyday prose" (RQ160, f, 23, Ques). Others defined quotations in terms of what they are not, especially "cliché" (RQ148). This suggests that quoters do indeed pay attention to quotations in part because they stand out as particularly powerful or charming linguistic artifacts.

Related to this, quoters (10 reader-quoters, 3 bloggers) noted that quotes need to be short.<sup>39</sup> Quotes are "not too long" (RQ28, Ques) and "short and concise" (Blogger81, f, 26, Ques), and they express a thought "succinctly and well." But this is not brevity for brevity's sake. A quote must pack a wealth of meaning and/or emotion, despite its diminutive nature, "condensing something that could take a paragraph into a few lines" (RQ89, f, 28, Ques). And a quotation must not be *too* short. Another quoter (RQ73, f, 33, Ques) had a hard and fast rule: "no more than four lines of text." One quoter (RQ112, f, 23, Ques), a native Swedish speaker, had a term for this balancing act:

We have a word in Swedish, it's called "lagom" it means something is not too much, not too little, but... lagom. Just about the right amount.

Expanding upon her definition of *lagom*, this quoter brought up the fact that this word also describes the meaning of the quotation, which is neither too "deep" nor too "shallow." A good quotation, then, can "make you ponder for a while" without "gett[ing] stuck on it for a longer period of time." Her description seems to suggest a way of reading quotations that is neither "hyper" nor "deep" attention but somehow in between. Another quoter, however, linked the brevity of quotations to the medium of Tumblr; a quote must not "[go] so long that the person reading loses the thread or simply scrolls past it" (RQ93, f, 21, Ques). One cannot depend upon

---

<sup>38</sup> Recall that Figures 2A and 2B of Chapter 3, the dendrogram visualization of the top 100 tags used in the GQ Hashtag Sequences Corpus, did not exclusively consist of English hashtags. Tags such as "#amor," "#frase," and "#zitat," also made the list. This is yet more evidence that quoting in general on Tumblr is not a purely Anglophonic practice.

<sup>39</sup> For purposes of coding, I considered these comments all examples of how quotes must be "brief." Calling a quotation is "short" does not mean that it is linguistically powerful in the same way as calling it "succinct" or "pithy," words that combine description of *length* with something like *aesthetic power*. I chose to consider these "brief" quotes separately as brevity immediately has to do with attention. (Short quotes may require less attention.)

the extended attention of the reader.

**Semantic aspects, especially "resonance."** Quoters often pointed a constellation of non-linguistic qualities of quotations that make them suitable to be repeated and shared. In general, quotations were seen to have some kind of affective or cognitive effect on the reader, though these effects were described with varying levels of specificity. Several quoters (7 reader-quoters, 1 blogger) suggested that quotations must be "inspiring" in some way, pointing to a positive message. Quotations might also "make you think" (RQ87, f, 26, Ques) or more generally "evok[e] strong feeling (both positive and negative)" (RQ86, f, Ques).

However, for many quoters (33 reader-quoters, 9 bloggers), the emotional or cognitive effect of quotations is best described as one in which the quote "resonates" with them in a highly personal way. That word "resonate" emerged as a key word that respondents frequently used to describe what it is that a good quotation does vis-à-vis the reader (6 reader quoters used this word, and 1 blogger). Other quoters got at this sentiment in other, sometimes more complex ways. A quote must have "[r]elatability to my own life" (Blogger54, f, 21, Ques), describing "exactly what I'm going through" (RQ117, f, 21, Ques), and so a good quote is one that "matches to personal experience" (RQ155, f, 24, Ques). If a quote expresses some sentiment, it must be "something [the quoter] agree[s] with" (RQ102, f, 29, Ques). Quotations must "validat[e] an individual's thoughts and feelings." Quotes can also "resonate" with a thought and feeling inside quoters that they do not normally express. A good quote "touch[es] on the normally unspoken and intimate" (RQ133, f, 35, Ques).

This does not mean that the quoter is unique in "resonating" with a quote. Quoters drew attention to the ways that quotations must resonate broadly, "with you and other people" (Blogger20, Ques) who are from "all different kinds of background[s]" (RQ87, f, 26, Ques). A quote offers "[s]omething everyone can relate to in some way" (RQ56, f, 23, Ques), and thus it should be "universal", expressing a generalized "human" sentiment (RQ120, f, 27, Ques). Another quoter went farther to describe what characterizes this universality: a quote must be "inoffensive," presumably because an offensive quotation would alienate readers rather than "resonate" with them (RQ126, f, 24, Ques). Quotations should be applicable not simply to the lives of many different types of people but also to multiple situations" (RQ5, n, 26, Ques). Quotations are paradoxical: they must be "[u]niversal and yet personal at the same time" (Blogger76, f, 25, Ques).

The need to make quotations "universal" in this paradoxical way specifies a certain relationship between the quotation itself and the text in which it originally appears. A quote must "make sense out of context" (RQ28, Ques) so that it "goes beyond the source text" (RQ79, 25, Ques). Other quoters offered a complex theorization of the relationship between quotes and their sources. Quotes can either be "powerful because of their original context and the weight of memory they evoke" or they can "stand powerfully on their own" (Blogger58, f, 40, Ques). This last quoter gave the example of a quotation from a book by Jeanette Winterson: "What is the measure of love lost?" This quote, she asserted, "can communicate effectively on its own." Another quotation, "I hope you weren't disappointed," makes sense only "when you know that it was what her biological mother texted her after their first meeting [as] adults." There is not one type of quotation, since a quote may be "something that is emotionally strong" or "important to the plot" (RQ59, f, 19, Ques). Thus, when speaking of the kind of attention that quotations solicit, it is necessary to distinguish not just between "hyper" and "deep" but rather between attention that is channeled back towards the book from which a quote came and attention that is

not. A good quote, according to quoters, expects the latter form of attention.

In this respect, quotations from popular sources may be different than quotations from unpopular sources. In the case of a quote from one of the *Harry Potter* books, "you can presume that most people would know something of the characters and context" (RQ93, f, 21, Ques). It is almost guaranteed to be "relatable for a large number of people" simply because most people are familiar with the story of Harry at Hogwarts. This was the only quoter who alluded to the way a good quotation can refer specifically to its story. However, even this respondent noted that quotations from less popular sources must resort to "general ambiguity" and "emotional depth." In this case, certain linguistic qualities of the quotation may assist in creating this "ambiguity." As this same quoter explained in the questionnaire:

Taking a quote on loss of a loved one, for example, is often more widely received when the gender-specific pronouns are neutral or not present. This way a person who lost their mother and a person who lost their younger brother can still connect with this same quote.

A "gender-specific" quotation, one about "he" or "she," would make this quotation less easily decontextualized and applied to various readers' relationships. This explanation gets at the way that the formal characteristics of quotations solicit certain kinds of attention from readers, modalities of attention that are inextricable from the purposes that underlie the practice of quoting itself. The following section considers these directly.

Before that, however, a final note: two quoters did mention the fact that the author of a quotation is important to the quotation's value. A good quote is one by "a famous person" (RQ104, f, 24, Ques) with "an associable name" (RQ11, f, 18, Ques)—presumably, a name with which one has pre-existing association. This hints at the possibility that factors external to the words of the quotation itself may also assist in capturing the attention of quoters, a possibility explored further in Chapter 5.

## Why Quote?

**Quoting is a form of self-expression.** In analyzing the questionnaire data (Appendix A, Question 3), one key pattern that emerged was quoting on Tumblr is a way of expressing their own feelings and emotions. Of the 157 respondents, 28 reader-quoters and 3 bloggers drew attention to the ways that quotations are self-expression. This purpose was sometimes entangled with the belief that quotations are somehow more able to express their emotions than they themselves are. Quoting is "a way of putting indescribable feelings into words." Not everyone is "well equipped with words" and so it is useful that "someone else has already put their feelings into words" (Blogger51, m, 25, Ques). Quotes are "more succinct" (RQ73, f, 33, Ques) and more "eloquent" (RQ75, f, 21, Ques) than quoters themselves could be. "[P]eople sometimes have trouble expressing themselves," (RQ35, f, 28, Ques) and quotations give them a way of doing so."

Such statements cast quotations (and, implicitly, their original authors) as more articulate, better with words, than those who quote. Yet, even if a quoter feels that she could not have written a quote, she nonetheless *appreciated* it. When a quoter discovers that "someone else has already put their feelings into words" (RQ31, f, 25, Ques), the implication is that emotions exist separately from words. One may feel emotions or experience certain thoughts or ideas without being able to articulate them. If a quoted writer has a unique way with words, the sentiments behind these quotations are not unique.

A quoter (Blogger78, Chat) who admitted to being "not a good writer" and "not as good with words" as the writers she quotes still had something positive and affirming to say about herself, as well about as people in general:

We are not poets yet feel what poets feel. The majority of humans are poets so far as emotions, experiences, mental states, etc. go but not all of us are writers haha  
The ability to feel emotions and to put them into words are two different things. The latter may be much rarer and beyond her ken, but she is a "poet" in her capacity to feel emotions and experience the world.

*Self-expression can be a form of self-care.* The fact that quotes *can* express one's emotions does not explain *why* quoters would turn to this form of self-expression. Perhaps quoting is just one of many such forms of identity performance, one that "informs [our] public image in the same way our manner of dress or speech does" (Blogger27, Ques). Quoters also pointed to some ways that quoting as a form of self-expression is a particularly important and powerful form of *self-care*. As one quoter put it, even quotes that thematize "depression or negativity" (rather than inspiration) may have a "positive emotional impact" by providing "solace" (RQ11, f, 18, Ques). Another major theme that emerged in coding was that quoting allows individuals to form imagined networks of individuals who feel the same way as they do. In short, quotation is a way of feeling "un-alienated" in one's emotions but rather is "connected" to others in the sense of feeling the same way.

21 reader-quoters, in addition to 3 bloggers, drew attention to the ways that the point of quotation is to recognize that one's emotions are not unique but are shared by others and are even perhaps universal. To see a quote that expresses one's emotion is to know that one is "not the only person to have had that thought or emotion" (Blogger76, f, 25, Ques) and is "not alone when going through certain aspects of life" (RQ80, f, 23, Ques). The quoter discovers an emotional affinity with the author of the quote; in one quoter's formulation, there is a special solace to be gained from knowing that an author from "a different era" shared the same emotions (RQ104, f, 24, Ques). Yet quoting in the networked space of Tumblr provides other vectors of connection:

I share quotes that have resonated with me, and I think it's interesting to see others react to those quotes or post their own. It underscores some kind of common thread between us, even though we're strangers. (RQ79, 25, Ques)

Notice that this form of connection is not direct *communication* but rather *observation*— "see[ing] others react" to a quote (presumably by "liking" or "reblogging" it). While one quoter did mention in an interview that she did actually make friends via Tumblr through chatting with fellow quoters, this is not mandatory. Another quoter described quoting as "a way of communicating without being required to have an actual conversation" (RQ34, f, 24, Ques). She continued:

I can reblog a quote or someone can reblog a quote from me and by doing so, there is a sense of understanding and, sometimes maybe even, agreement among the bloggers involved.

Reblogging is a latent symbol of agreement, even shared feeling. Because of this, quotation can "give a sense of community" (RQ50, f, 24, Ques). Along with helping people to feel "less lonely" (RQ88, f, 20, Ques), this can simply afford the quoter "validation" (RQ160, f, 23, Ques); if others have felt one's emotions, there must be nothing wrong with feeling them.

As one quoter explained, quotations are paradoxical: they "reveal something" personal

but "in a non-personal way" (Blogger91, f, 49, Ques). Even if quoters *could* put their emotions into their own words, using quotations as a way of combating alienation has its advantages. "[S]traight up saying how you feel" is "more personal to you" but also "much harder for someone else to connect with." (Blogger57, f, 23, Ques). It is in their lack of specificity to any one person—the fact that they convey personally-held feeling but are not *too* personal—that allows for various people to feel through them, forging "connection" or community. Another advantage of this paradoxical quality is that it is somewhat less vulnerable than saying what one has to say:

There's also maybe a built-in barrier because if someone asks why you posted a quote, why you feel that way, you can say, "It's just a book." (RQ35, f, 28, Ques)

Another quoter expressed a similar sentiment:

Reblogging quotes is a subtle way to show what you're thinking or how you're feeling without having to actually be accountable for what's being said. (Blogger54, f, 21, Ques)

Expressing oneself through the words of others gives the quoter a "rhetorical shield" since they can "distance themselves from [the quotation's] sentiment if or when challenged" (RQ35, f, 28, Chat).

Tumblr in particular may be an important space for this sort of quotation, as quoters often spoke about what makes this network in particular a space for forging imagined networks of affinity. Tumblr is "a social channel made to be more emotional and raw than...Twitter or Facebook" and Tumblr users are "more receptive to posts that make them feel something" (Blogger57, f, 23, Ques). Tumblr is "anonymous" compared to other social network sites because you can interact primarily with "strangers" (RQ88, f, 20, Ques). (The "anonymity" also likely derives from the fact that Tumblr users often use pseudonyms.) Another quoter observed that on her Tumblr blogs she can share "a side of me that my own real friends [do] not usually know" (RQ100, f, 20, Chat). Likely it is because of this relative anonymity and possible separation from one's "offline" relationships that Tumblr is "a safe haven when it comes to expression of people's most inner thoughts," (RQ139, f, 20, Ques) including through quotations. Quoting in particular is a way for users "[t]o express themselves and not feel judged" because "most people who will see what you post don't know you in person" (Blogger43, f, 18, Ques). There is safety amongst strangers.

*"Inspiring" quotations and other-care.* As self-expression, quotations might express a feeling that are negative (e.g. melancholy, anger), since sharing this emotion, and recognizing it in others, can allay feelings of aloneness. Though less common than the self-expression function, some respondents (9 reader-quoters, 2 bloggers) did mention in the questionnaire that a main reason for sharing quotations is that they contain some positive content (for instance, that they can be inspiring or motivating). This trend suggests a more direct way in which quotations are tools of self-care: they contain some positive message that in some way benefits or perhaps teaches the reader. The circulation of such quotations can be a form not just self-care but also of other-care.

RQ111 (f, 25, Ques) cast the power of quotations as the power to "enlighten someone" and, in the process, even "save a life." Wondering if this was somehow hyperbole or metaphor, I asked her about it in an interview, during which she confirmed this starkly unironic faith in the enlightenment offered by words: "the right quote can hit [someone] hard and allow them to realize life is worth living." Other respondents were less extreme in their evaluation of a quotation's salutary power. Wrote Blogger146 (f, 26, Ques):

Reading often inspires and provoke thoughts and emotions from people. I think when

people find quotes from things they've read or find a passage that is particularly inspiring or uplifting, they want others to feel the same way and so they share quotations.

Note that both of these respondents' explanations are related to but different than those related to quotation as a way of forging an imagined affective community. In these descriptions, the quote is "inspiring" or "enlightening." Another word for "enlightening" might be "pedagogical"; they do not just *reflecting* one's thoughts but actually improving these thoughts by casting brief but necessary light upon one's life and the problems within it.

These explanations of the logic of quoting figure the quotation as a morsel of truth or wisdom that has the potential to improve those who read it. Another Tumblr user, one who identified as a relogger rather than a reader/quoter, evokes a different concept to explain how the quotations she reblogs might help others:

I think God can work through what I share and get it to someone who needs it. (RQ144, Chat)

Certainly the helpfulness of quotations can be explained without invoking a deity, yet this quoter's rather theological understanding draws attention to the complexity of systems like Tumblr. Quotations are circulated according to the actions of the users who blog and reblog them and, to a lesser extent, by the algorithms that may recommend or highlight certain accounts or users. Yet this non-divine circulation is also a mysterious process, since (as has been mentioned) one does not necessarily know the reblogging user and certainly does not know this user's motivations for reblogging. Posts on Tumblr do indeed "move in mysterious ways." This blogger's imagination of the network itself as a beneficent if unknowable force underscores the fact that when one cares for another via quotation, this "other" is not so much a specific person as a person *out there*.

**Quotation is a gendered literacy practice.** Lingering behind this analysis is the gender of quoters. Is there something about quotation as a practice of self-care and self-expression that makes it particularly "feminine"?

Quoters believed this to be the case.

Certainly there is a danger in essentializing women in any way as "emotional" and prone to collective sharing of emotion, in contradistinction to the stolid and emotionally reserved (if not emotionally repressed) male. In fact, the quoters I asked about this issue generally suggested that there is some truth in such essentialism, even as they sometimes occasionally registered their discomfort with it. One quoter argued that women/girls are simply more "empathetic" than men/boys, but not without a bit of nervous laughter:

I know it's like unpopular now to say that women are more empathetic and what not than men are but it's just kind of true haha (Blogger98, f, 18, Chat)

One of the rare male quoters seemed perplexed by the gender-imbalance but did wonder if quotation is particularly feminine because it is a form of "self-expression":

Do men post private details or reveal emotions as much as women do? I'm not sayin this to reinforce gender stereotypes. [sic] (RQ83, m, 27, Chat)

Another quoter (RQ100, f, 20, Chat), a college student herself, brought up the notion of "stereotypes" as concepts that mandate or reinforce certain behavior:

Because girls are stereotyped (or expected) to be expressive, we tend to express what we think and feel freely. And for boys, since they are considered to not show that much emotions or else, they might get called 'pussy' or 'a feminine person' or 'gay'...

She suggested that it may simply be "safer" for male Tumblr users to not quote, figuring them as



repressed by society. However, another quoter (RQ95, f, 25, Chat) linked book quoting not simply to female expressivity but to the fact that any expression by girls or women occurs in the context of social forces that work to silence them. According to this quoter, the speech of women and girls is still regularly policed, critiqued, and dismissed as either trivial ("that's nice dear") or outright attacked ("typical dumbass feminazi!"). In this context, women like her are "forced into anonymous blogging to get thoughts out." Seen in the light of Tumblr as a whole, the anonymity provided by quotation as a form of self-expression may be yet another layer in this gendered self-defense. RQ35 (f, 28, Chat) agreed with this idea while also complicating it: for her, quotation is of particular use not just to women/girls but "other minorities" as well. As she explained, "oppression has a lot to do with being silenced by those in power." Quoting is a way of speaking while under the pressure of this enforced silence.

**Quoting and books.** The reasons for quoting specified above are interesting in part because they are not really about *books* as such. (They may apply just as well to quotations from songs or TV shows.) Still 13 book quoters and 1 blogger explicitly described a reason for quoting that had something to do with books and bookish culture.

*Quotations and bibliophilia.* Quoting can be a part of engaging in "bibliophilic" culture on Tumblr. The practice of posting "#book quotes" on Tumblr is also a way to participate in a general "bibliophilia." This may mean finding "others who like the same books" as the quoter, in effect "[b]onding over favorite literary works" (RQ128, f, 30, Ques). Or, as in the case of one quoter, simply finding those who love books in general:

In day-to-day life it is rare to find avid readers, so it is refreshing and inspirational to come to a website like Tumblr...where you can be passionate about literature and encounter others who feel the same way. (RQ120, f, 27, Ques)

As seen in the previous chapter, the existence of tags like "#booklr" suggests that Tumblr is a home for dedicated and passionate bibliophiles. The words of these two users confirm that this is an important aspect of Tumblr's quotation culture.

*Quotations as advertisements.* Quoters view their quotational activity on Tumblr as a form of "promotion" for particular books as well as other media:

[Quoting] helps promote your favorite authors, poets and speakers. On a website like tumblr, where users connect to one another through shared media interests, this is a great way of encouraging people to become part of a fanbase (RQ7, f, 25, Ques)

Quotations can "inspire someone to explore the work I'm quoting" (RQ23, f, 25, Ques) and so to quote is to "recommend reading material" (RQ86, f, 30, Ques). One quoter spoke of this as a way of "supporting their fandom" (RQ134, f, 24, Ques). Tumblr users are conscious of the way that quotations are essentially advertisements.

But do the advertisements work? One quoter described the discovery of books as a primary reason for participating in quotation culture:

I...find myself jotting down the titles of books or the names of authors I haven't read. Finding a quote that I love is a great way to find something new to read (RQ34, f, 24, Ques)

Others too described adding to their "to-read lists" (RQ102, f, 29, Ques). At other times the connection between a quote and a longer text does not even require such readerly labor. According to one quoter, "[s]ometimes the quote is from an online piece with a link to it so i

follow the link and read...there" (RQ104, f, 24, Ques).

In response to a subsequent question that directly asked respondents whether they had ever discovered a book through Tumblr quotes, only 2 of the 93 book quoters who answered this question responded "never." The most frequent answers were that they "sometimes" (50) or "frequently" (25) had discovered books through quotations. Quoters reported discovering very different types of books through Tumblr. Eight mentioned the works of John Green, echoing his general popularity on Tumblr and in book quotes in particular. But seven described reading the works of Dante. Four discovered the poetry of Rupi Kapur, one of a cadre of young poets who publish their works exclusively or at least first on social networks (in her case, Tumblr and Instagram). One philosophically-inclined quoter used quotations to discover both Roland Barthes and Alain Badiou, though most of the titles given by respondents could be classified as literature, mostly fiction. Quotations can lead Tumblr users to discover authors who become extremely important to them. Wrote one Tumblr user:

I was never really interested in this author until I saw some of the quotes from his works on tumblr, and now he's one of my favorite authors. (Blogger109, f, 19, Ques)

Wrote a reader-quoter:

I'd never really taken the time to check out any of Wilde's works before I saw some of the most beautifully worded quotes on tumblr from this book. He's become one of my very favourite authors. (RQ134, f, 24, Ques)

Quotations are of course not the only way through which readers may discover authors, but perhaps they have advantages over other methods. One quoter I interviewed described her process of posting the first line of whatever book she happens to be reading. "A book synopsis only says so much," one quoter (RQ86, f, 30, Chat) explained, "but the first line gives you a sense of the author's style and the book's tone." One quoter discovered a novel after seeing "lots of quotes" but also "memes" and "fanart" about that book. In such cases quotation must be seen as but one aspect of a larger array of "fannish" texts.

While 93 book quoters described at least "rarely" discovering a book through Tumblr quotes, only seven reader-quoters (in addition to one blogger) described either finding books or recommending books to other readers on the network as a primary reason that quoters quote. This finding suggests that the discovery function of quotation, though not a primary or conscious reason most quoters quote, is a significant epiphenomenon or side-effect of the primary purposes of quoting—those related to self-expression and self-care. Still, quotations are clearly an important tool for luring readers to books, encouraging and supporting off-line, print-based practices of "deeply" attentive reading.

*A note on corporate activity.* Quoters are most often individual readers, their quotation seen as an act of literacy that is both connected to and separate from the corporate-produced texts that they excerpt and illustrate. However, it would be somewhat incomplete to envisage Tumblr as purely a space for the cultivation of bookish "affinity groups." Already this discussion has alluded to the fact that not just books-lovers but also authors themselves who participate on the site (e.g. John Green). Still, it came as a surprise when one of my questionnaire's respondents described itself not as a person but as a business—a literary agency. The operator of this Tumblr account, an employee of the agency, wrote that they quote "excellent lines" from "books that our clients publish" (RQ69, f, 21, Ques). In a brief interview, this representative suggested that while "photo content, videos and links tend to drive more interaction," it is true that "the occasional fantastic quote does well." This person also suggested that quoting on Twitter is more likely to

spark a discussion about a book than would quoting on Tumblr. Obviously this agency has a different goal than most quoters: to create "engagement," directing attention toward *specific* books by its clients.<sup>40</sup>

Although the focus of this study of individual reader-quoters, it is clear that the official organs of book culture are at least marginally involved in quotation and in bibliophilic culture on Tumblr more broadly. Out of the five most prominent publishers of fiction in the U.S. (Penguin/Random House, HarperCollins, Simon & Schuster, Hachette, and MacMillan), each of them has at least one official Tumblr account, either for itself or for one of its imprints. The Tumblr blogs managed by publishers have posted book reviews, pictures of cover art, and quotations from books that they have published. These publishers seem to use quotation as a form of advertisement for their own books, yet they may also reblog quotes from other Tumblr users and post quotes by authors whom they do not themselves publish. Such actions blur the boundary between the producer of the book and the consumer. Social media in this case a multi-national enterprise, employing thousands of workers globally, may shrink itself down to the human scale, becoming "just another blogger."

Is this a successful gambit? The present research design cannot shed light on whether or not publishers' and agencies' participation on Tumblr and specifically their quotation practices drive sales. Still, the presence of these accounts is itself a good reminder that Tumblr is not a utopia separated from the markets and institutions that regard their activities on the site as another opportunity to convert "user engagement" into literary commerce.

**John Green's "shtick."** As shown in the previous chapter, the Young Adult author John Green is the most popular author quoted with the "#book quote" or "#book quotes" tag and among the most frequently quoted authors on Tumblr overall. This may be for any number of reasons, from the fact that his books have been made into major motion pictures or that he has long maintained an active web-presence on Tumblr and on other networks. However, his popularity could also be attributed to the ways that his books—or at least quotations from them—fulfill some of the major purposes of quotations discussed above. Quotes from books by John Green have "an uncanny way of putting into words thoughts and emotions that many people find difficult to describe" (RQ7, f, 25, Chat).

One quoter (RQ75, f, 21, Chat) remembered discovering John Green on Tumblr when she was in her mid teens. She pointed to the following extended passage as one of her favorite John Green quotes:

I wanted so badly to lie down next to her on the couch, to wrap my arms around her and sleep. Not fuck, like in those movies. Not even have sex. Just sleep together in the most innocent sense of the phrase. But I lacked the courage and she had a boyfriend and I was gawky and she was gorgeous and I was hopelessly boring and she was endlessly fascinating. So I walked back to my room and collapsed on the bottom bunk, thinking that if people were rain, I was drizzle and she was hurricane.

And she offered this explanation of why these words stood out:

This was my favorite quote back then. and i think the reason why me and so many other teens liked this or related to it is because adults tend to think of teenagers as hormonal or

---

<sup>40</sup>"Engagement," as Green and Jenkins (2009) observe, has become the buzzword of enterprises who seek to capitalize upon the devoted and engaged social media activity around corporate media in the context of online participatory culture.

"sex driven" while we just wanted someone to understand. someone to talk to and have a mental/physical connection with. also i think as teens we all feel gawky, boring, out of place, awkward. he just managed to put those feelings on paper and make them sound... smart, eloquent.

John Green's quote expresses the true feelings of this teenage quoter. Unlike other "adults," who misjudge teenagers, he was able to understand her and speak for her, while also making these thoughts themselves sound "eloquent." It is not just that the quote puts inexpressible thoughts into words but also that, by doing so in an artful way, it validates these emotions themselves. Adults are clueless, but teenagers hold inside them thoughts and feelings—valid ones, "smart" ones. Likewise, RQ7 (f, 25, Chat) suggested that John Green's quotes are so often quoted because his books themselves "validate" the emotions of Tumblr users, especially teenage ones:

I think it's no mistake on John Green's part that his writing tends to be either in first person or his most quoted lines within dialogue. Younger people who read it can envision themselves speaking like that, thinking like that, and being spoken to like that. I guess in a way, while his plots are typically relatively simple, that emotional, poetic language speaks teenage readers while remaining accessible. The writing doesn't talk down to them and it shows typically YA characters as intellectual and beautiful - it validates teens as a whole, and as younger people make up the majority of this websites users, it's unsurprising that tumblr blogs like to reference his work.

This theory of John Green's popularity rings true with the description of RQ75 (f, 21) that John Green's quotes made her teenage self feel "smart": as a whole, his books allow teenage readers to "envision" certain characters' words as their own while also implicitly complimenting (or, to use a less kind word, flattering) these readers as "intellectual and beautiful." She further suggests that Green's books are linguistically structured (through first-person narration and dialogue as well as "poetic" but "accessible" language) to be quotable. This "accessibility" perhaps explains the popularity not just of Green but also other authors of YA fiction and children's literature::

Authors like Chbosky, Jennifer Niven and Benjamin Alire Saenz get quoted a lot for the same reason. It's beautiful, but easy to understand too.

During an interview, I asked RQ7 (f, 25) whether it would be fair to call Green's quotes "emo"<sup>41</sup> She replied that this term is "a fair word for" for the work of John Green:

It's angsty, but angst can kind of feel empowering for the same reasons. I think most teenagers on some level go through a dramatic, emo-y phase which might feel a bit silly as you get older, maybe, but when you're young it's all encompassing and you're naturally drawn to media that mirrors it

For her, John Green's novels about emotional but intelligent teenagers are designed to "mirror" the perhaps overwrought experiences of young adulthood. Still, it is interesting to note that, even in the case of RQ75 (f, 21), who originally identified so strongly with that extended quote about sexuality and desire, her identity is in some sense at odds with the identity of the character, at least in one very basic way: the main character is a teenage *boy*. This quoter noted that "growing up the books i related to most and identified with usually followed a male protagonist," though she attributed this to the fact that there were not "a lot of options on the market." Even in her reading of the original quote, she "decontextualized" the quote at least somewhat, understanding it as a general sentiment about general desire (rather than a *male* sentiment about *male* desire).

The description of John Greene as "accessible" because it is in the first person (RQ7, f,

---

<sup>41</sup>As in "emotion." The word referred originally to a sub-genre of angsty punk rock.

25, Chat) raises the possibility that Green's prose itself is "quotable" (perhaps even by design). In the questionnaire, another quoter who discovered Green's work through quotes paid the author a rather backhanded compliment:

...John Green novels...not a fan of his general shtick but he can often have a wonderful way with words. (RQ95, f, 25, Chat)

In a follow-up interview, this reader further specified exactly the nature of this "schtick." To her, John Green's writing is characterized by "relatable blandness" that is "peppered with tiny gems":

To me, John Green is watered down fiction. He does produce lines and passages that are good and are very relatable, which is why I think they get passed around so much. But he feels like a few deep passages amidst a sea of shallow lines. When reading his work, it leaves you wanting more, but I wouldn't say more of the story. My other writer and bibliophilic friends have discussed this and came to the conclusion that what we were left wanting was something richer. Better vocabulary, stronger details, less predictability

This is only one quoter's opinion, yet it raises the interesting possibility that certain prose styles are in fact designed for quoting; narratively shallow and predictable, they nonetheless contain a few "deep passages" that warrant sharing.

But this is only one theory of John Green's quotability. RQ26 (n, 21, Chat) drew a distinction between John Green's quotes and his books. She noted that she "has seen posts on Tumblr accusing him of only writing about white, privileged characters." Launching a similar criticism, RQ75 (f, 21, Ques) noted with disdain that all of Green's books tend to include a sweet but nerdy male main character and an unrealistic "manic pixie dream girl." Still, according to RQ26 (n, 21) "the way he words things can be highly relatable, particularly for those with mental health issue[s]." She continued by linking John Green's popularity not to a teenage audience but an adult, bookish audience:

I think the reason he is so popular on Tumblr is that the stereotypical book blogger is an introverted, anxiety or depression sufferer, and it is sometimes easier to express how one feels behind the mask of another person's words than acknowledging and describing personal, painful emotion.

Her words point to the power of Green's quotations to voice one's emotions while also shielding them. Yet, for her, the audience of these quotations are a certain kind of adult—"book bloggers" who suffer from one or another psychological malady. And so the books themselves, which deal so directly with the everyday lives of teenagers, are less useful than the quotes from them. Seen as works in their entirety, novels like *The Fault in our Stars* are seen as politically suspect for their focus on white characters: still, according to this quoter, "the way he words things can be highly relatable, particularly for those suffering with mental health issues."

These quoter's theories of John Green in some sense contradict each other, but they need not be seen as mutually exclusive. Perhaps teenage and adult readers appreciate his work as well as quotes from his work in different ways. In general, however, these theories help explain Green's popularity by drawing attention to a consistent characteristic of those quotations from his work that have become so popular: they express what the quoter herself feels, in the process "validating" these emotions.

### **Modalities of Attention**

Quoting as an embodied practice involves decisions about how to actually manage one's quotes—how to physically extract them, and whether to post them in the process of reading, or

after reading. Different quoters explained variable and idiosyncratic practices of quoting.

One strategy is to simply post quotes while in the flow of reading. RQ95 (f, 25, Chat) described her process of quoting a particular novel in this way:

...I was just posting things as I went. I had [the book] in front of me, something resonated with me, and I attempted to use the magical interface of Tumblr Mobile to log them.

Her words casually gesture towards a way of reading that would seem strange, even anathema, to those who feel one should "get lost" in the novel. (As this quoter admitted, she is a "multitasker.") If it is a truism of contemporary culture that the smartphone is the omnipresent distraction, the pest that one can never quite put down, this can be true even in this scene of reading. Yet what she describes also is an interesting reversal of the smart phone as a technology of distraction. The "ring" that disturbs the reader comes from within this reader: when something "resonated," she was cued to turn from the page to the screen.

Other quoters may be less comfortable with this form of reading-while-blogging: Another quoter (RQ72, f, 19, Ques) explained that she used "read and blog at the same time" but ended up finding this way of doing it too "annoying." These days she merely marks down the page number of a quote; then, after reading, she goes through her document and looks up the quotes before adding them to Tumblr. This new practice comes with an added benefit: "[I]t's like you're reading the book again and that's really nice." In some ways, for this quoter, quoting on Tumblr may also provide a *deeper* reading experience than reading without Tumblr.

In fact, RQ83 (m, 27, Chat) made exactly this point as well. During a follow-up chat, he recalled a time when he was younger when he would "[swallow] books whole," becoming fully immersed in their worlds. For this quoter, however, such a way of reading is not actually "deep"; after devouring a book, he would not be able to remember any of the "finer details." Posting quotes to Tumblr, for this quoter, is a way to actually make the reading process slower, more reflective. As he describes himself, he is a "thoughtful reader" who "enjoys sentence structure and mechanics." For him, quoting is less a matter of social media *per se* and more an opportunity to slow down and savor language, a "deeper kind of reading" that "allows for greater appreciation of the work." This quoter seems a connoisseur of language; for him, a quotation is an artifact of beautiful writing to be collected, studied, and appreciated.

Likewise, RQ100 (f, 20, Chat) described the process of quoting books while reading as a kind of distraction:

NOT a bad distraction however, it is a distraction wherein I just found a beautifully written sentence/paragraph and that I know I could share it to others by posting in social media sites. So I would take time writing it or like spend a couple more minutes trying to 'feel' the words and understand it deeply within me.

For her, quoting is an opportunity to slow down the reading process—to "spend a couple more minutes" with the quoted words. Even rewriting the quote for a Tumblr audience (a process that "takes time") forces this slower, dilated reading experience.

Yet even those who post quotes from their own reading may also enjoy becoming lost in a deluge of micro-texts. Another quoter alluded to the ways that she consumes quotes in the course of describing Tumblr's quotation culture in general:

While novels and poetry can provide a similar release, they're also more of an investment. Of time and concentrated effort, in particular. Quotes on Tumblr are generally brief and broken up with other media—comics, memes, cute puppies, pictures of things the blogger is interested in, like horses or nature or fandom-related posts. It's easier

to read, like, and reblog a few quotes during your lunch break or while working on your phone or laptop...I would definitely say that quoting is a means of breaking up a person's day with something calming or uplifting. I personally browse inspiring quotes between tasks, to take a breather of sorts. Tumblr's especially useful for this because of the mixed content and the ease of just scrolling and choosing what to pay longer attention to and what to give only a brief glance to. (RQ93, f, 21, Chat)

This quoter makes an argument for the value and necessity of "hyper-attention." In a most literal sense, quotes from novels and poems take much less time to read than entire novels and poems, and so they can be "interruptions" in the day, especially the work day. One can "reblog a few quotes" during official times of non-work, like lunch, but also "while working on your phone or laptop," stealing a few moments of leisure. This quoter describes the importance of quotations being interspersed in what she calls "mixed content," which may also contain "memes, cute puppies" and other fare. No one text, or even one type of text, gets the most attention. It is as if each text is a distraction from the previous one—a series of textual *amuse-bouches* that delight in part through their differences. This Tumblr interface itself, an endlessly scrollable feed of diverse media operated with a single click of the mouse or flick of the screen, seems designed for this sort of consumption. For this quoter, it would be wrong to use the phrase "reading a quotation." The multiplicity of quotations is part of what makes them appealing, allowing for a somewhat passive, somewhat active interstitial experience of peace (a "breather") in which the quoter does not so much seek out a single piece of wisdom or inspiration but grazes from a buffet of these qualities. Quotation, along with other Tumblr artifacts, does amount to a distraction, but this can be an intentional distraction—a mini-meditation.

Other quoters seemed more ambivalent about the way that they consume quotations and media on Tumblr especially. BQ16, for instance, described the rush that he feels sometimes when scrolling through his feed:

I don't know about everyone else, but I move at a pretty fast clip, my eyes scanning rapidly to see if there is something of interest. I'm highlighting the speed here because I feel like we're all always in a rush and never standing still, not even if we're flicking through our feeds on tumblr. (RQ93, f, 21, Chat)

It is remarkable that this is the same quoter who felt that *posting* his own quotes actually slows down the process of reading, allowing him to appreciate texts more. His words point to the ways that even for a single individual participating in quotation culture may mean engaging in different forms of attention, some of which feel more satisfying than others. For him at least, the difference seems to be one between *producing* quotes (from books) and *consuming* them (on Tumblr itself).

Another quoter (RQ131, f, 25, Chat) also drew attention to Tumblr's as a site of "hyper" attention while also comparing it to the design of Goodreads. Unlike Goodreads, with its carefully categorized reviews and lists of books as well as quotes, Tumblr is much more chaotic. But the "randomness of your timeline" on Tumblr has a distinct advantage: it "creates the perfect opportunity to find quotes to like, hate and love." As an endless feed, Tumblr is designed for serendipity, but also for passivity: "In goodreads you find books, but in tumblr the books and quotes find you." Still, this reader does not eschew Goodreads totally but rather uses it in tandem with Tumblr, capitalizing on the strength of each in ways that these technologies' designers would not have predicted. For instance, she might encounter a quote on Tumblr and then visit its reviews on Goodreads. If the reviews are positive, she might buy the book, leaving reviews on Goodreads and quotes on Tumblr in the hopes that another reader might find it.

However, other quoters seem to regard Tumblr differently, casting it as a tool not just for *discovery* but also *organization* of quotes. Tumblr is a sort of "personal notebook" for quotes, a place where one can "keep track of quotes that inspire me" (RQ128, f, 30, Ques) One self-described "extensive" reader noted that "it is easier to keep all my favourite lines neatly organized on tumblr," making it a "tidy scrapbook" (RQ52, f, 24, Ques). Tumblr becomes then a "another way to preserve something you enjoy so you can easily come back to it" (RQ154, f, 26, Ques). This is in fact the *opposite* of watching a quote flow by in one's feed: the intentional archiving of quotations so that one may return to them later on. Certainly this could be considered a form of "deep" reading, since the goal is to dilate the experience of the source text over time.

These quoters describe a way of using Tumblr as a sort of personalized "commonplace book" (or, to use a more contemporary comparison) a "content management system." Even though Tumblr is ostensibly a "social network," quoters use it for private purposes—yet sometimes the social and the personal ways of using the site begin to bleed into each other. BQ7, describing the process of extracting quotations from a novel, noted that the fact that she shared quotations was somewhat of an epiphenomenon of a desire to record for herself important passages:

I wanted to save [certain passages]. I wanted to look at it later. And when one of my actual friends reposted it, I would wonder if they felt the same. (RQ95, f, 25, Chat)  
For RQ7 (f, 25, Chat), Tumblr's usefulness as a site for storing quotations has only part to do with its "archival" nature:

I used to have a notebook full of quotes from the books I liked, but I kind of preferred Tumblr. It's easier to find the quotes with this website because they are archived. And the atmosphere in here is nice somehow. When I post quotes about a book, sometimes I have people coming to me like "omg I see you're reading 'insert book here'! Are you liking it? Isn't it great?" and you can start amazing friendships.

The goals of making a personal archive are not incompatible with the more social aspects of quoting, including those related to emotional affinity ("wondered if they felt the same") and bibliophilic camaraderie ("omg I see you're reading...").

Another Tumblr user, one who described herself mostly as a reblogger of book quotes, likewise envisioned her act of quoting as a sort of proleptic activity stemming from a desire to encounter certain excerpts at a later point:

I use tumblr because it will pop up again over time whereas in a diary once I wrote it once it might not come back again. (Blogger122, f, 18, Chat)

In further conversation, it became clear that quotes will "pop up again" over time because she uses a feature that is specific to Tumblr: *queuing*. While the default method of posting to Tumblr is doing so in real time, bloggers can also add them to a queue that periodically publishes posts at specified intervals. As Tumblr's documentation suggests, this is "an easy way to keep your blog active and consistent," even when not logged-on. Her queue is a way to remind herself of a particular text at some point in the future. The queue itself becomes a "bot," an algorithmic agent which can be depended upon to "remember" a quote and to remind her.

Certainly I have only managed to adumbrate a few of the ways that the body, technology (particular devices), and cognition (including attention but also affect), intersect in the posting and reading of quotations. The larger point to be made is that even these few examples suggest that the simple binary of "hyper" vs. "deep" attention, while in some ways helpful, is not complicated enough to capture the diverse modalities of attention at work in quotation culture.



Between the isolated consumer of novels, lost in the book and transported to a world of her own, and the frantic social media user, torn between a thousand short texts, are these *real* readers, who shape their own experiences of literacy that are more textured and complicated than either of these two caricatures. They have their own preferences and ways paying attention to quotes. In fact, there is the possibility that individuals experience quotations in different ways at different times; the same quoter who deeply ponders the quote he or she posts on Tumblr may still enjoy the hyper-attentive experience of consuming quotes in the Tumblr feed.

**Connection to composition.** Quotation can be connected not just to other forms of textual consumption (such as reading) but also textual production (i.e. writing). Quoters rely on quotations to say what they cannot say (but nonetheless feel). Still, RQ35 (f, 28, Chat) suggested that quotations could be a "stepping stone" in a larger development as a reader and a writer. She suggested that quoters may begin by "figuring out how you feel" before moving on to "expressing that in your own words." Perhaps this is an autobiographical reflection, since this quoter herself identified as a writer. Her explicit formulation of quotation as a "stepping stone," however, was unique; most quoters seemed humble, even self-deprecating about their own abilities to write, despite a clear appreciation for the power of written language.

Still, writing and quotation can be entangled in other, less linear ways. In interviews, two different Tumblr users brought up moments in which quotations became the raw material for creative production. One respondent, who writes science fiction, suggested the ways that browsing Tumblr can help her get out of a creative rut:

Sometimes when I'm lacking inspiration, I search for quotes on tumblr, and suddenly I know exactly what (and how) to write....For example, I have a character that has this gloomy personality. Suddenly I find an Edgar Allan Poe's quotation that goes exactly with her time of life. Or maybe the language itself gives me inspiration with words when words escape my mind :) (Blogger92, f, 23, Chat)

She then described two different ways of using a quote. A character in her novel might quote Poe directly, citing him, or might simply utter something in his macabre, doleful style.

RQ110 (f, 32, Chat) described the role that quotation (and not just *book* quotes) can play in the production of media within fan communities. According to her, another fan of the TV show *Revolution* asked her for quotes to use in a new piece of fan-art. She obliged, noting that her comrade-in-fandom made use of the quote though unfamiliar with its source. Another quoter connected quoting to a more academic form of textual production:

Personally, I jot down quotes on Tumblr that move me or inspire me as I am studying for a masters in creative writing. (RQ50, f, 24, Ques)

In a brief follow-up interview, this same quoter explained in greater detail exactly how she uses quotations as "inspiration" for the kind of creative writing she is studying, playwriting: ...when I'm writing a new character or a piece of dialogue and need a certain emotion for it, I will revisit a quote that made me feel that way and try to recreate that emotion in a different context. It's not a matter of plagiarism because I rarely use the same wording but a means of welling up a particular emotion to be used. It's comparable to how my actors will use music to shift their mood depending on the piece they're working on. I use words to shift my mindset and work from there.

Unlike Blogger92 (f, 23), who described using the actual language of quotation, this quoter made it clear that "inspiration" does not for her mean the clever reappropriation of linguistic material (what she would call "plagiarism," what others might call "remixing"). Instead, the language of

quotations serve to trigger emotion within the writer; "inspiration" here means something like the production of intense feeling—or even a *reproduction*, since she can seek out previously-encountered quotation. The quotation can be an external storage device not so much of information as affect.

**Quotations and memory.** This dissertation has so far limited its gaze to "digital" media (Tumblr) and "old" media (especially books). But there is one other realm that is important to consider when discussing quotations. Five different reader-quoters mentioned that a good quotation should be *memorable*. This hints at yet another way that a quotation is not necessarily meant for quick, absent-minded consumption but rather integrated into one's memory. Quotations can exist in the mind just as they can on paper.

But what use is it to have a quotation in one's memory? In response to a pair of questions that asked quoters to remember as best they could a quotation that struck them as "especially powerful" (not necessarily one that they found on Tumblr) and then to explain what made this quote so powerful, quoters revealed some of the ways in which it is useful to have quotations memorized.

Memorized quotations can provide inspiration in times of great struggle. One quoter, recently "kicked out of the house," recalled this quote (unattributed, but it is from Tolstoy):

We are all in paradise, but refuse to see it.

She gave the following understanding of this quote's message:

I've been going through a rough time, being kicked out of the house, becoming homeless, working 3 jobs while going to school, and it honestly just makes me feel down, but I should be happy because I have a roof over my head, and I can provide for myself.

(RQ135, f, 21, Ques)

Quotations, waiting in one's memory, can also be mobilized during a certain difficult event. RQ35 (f, 28, Ques) remembered this quote from Dodie Smith's novel *I Capture the Castle*:

Noble deeds and hot baths are the best cures for depression.

She remembers this quote from time to time "when bigger things don't seem to be working out," taking it as a reminder that you need both big and small things to be happy." Another quoter recalled these lines, attributed to Charles Bukowski:

[A]s the living do not arrive, as the dead do not leave. I won't blame you instead i will remember the kisses our lips raw with love and how you gave me everything you had and how i offered what was left of me.

These lines also seem to play an enduring role in this quoter's emotional life:

I think it sticks with me because apart of me feels broken and I could never love someone as much as I could be loved in returned. And it's just a random quote I think about maybe when I'm fighting with my partner or feeling overwhelmed with certain emotions. (RQ39, f, 31, Ques)

Having the quotation in her memory, she can turn to it when certain emotions or recurring situations threaten to overwhelm her. Another quoter, remembering a rather inspiring quote attributed to Nabokov ("The breaking of a wave cannot explain the whole ocean") provided the following explanation of its meaning:

It reminds me that little mistakes, flaws, or faults do not dictate who I am as a person or anyone else for that matter. I think about it whenever I feel a little down about myself. (RQ47, f, 20, Ques)

This notion of the quote as a "reminder" was a common thread in quoters' descriptions of recollected quotes. A quote can "remind" a quoter "that I'm not perfect" (RQ112, f, 23, Ques), "how to live my life" (RQ147, f, 23, Ques), and "to avoid temptations" (Blogger137, Ques).

Sometimes, when the quoter is familiar with the book from which a quotation originates, the memorized quote in turn evokes the entire story. One quoter reproduced the following quote attributed to Sarah J. Maas:

She wore her scars like some women wear their finest jewelry.

This quoter (RQ145, f, 21, Ques) also noted the person (the "she") this quote describes:

Because Celaena has been through so much and to see that she doesn't shy away from it is very powerful. She isn't ashamed of her past and the scars it left. She uses it to make her stronger and to keep moving forward.

In this case, the quoter perceives the quote not as decontextualized wisdom but as a sort of linguistic totem that encapsulates the narrative tribulations suffered by "Celaena." Another Tumblr user (Blogger122, f, 18, Ques) remembered this extremely short quote from the final book in the *Harry Potter* series:

And all was well.

She wrote that she remembers this quote "because it's the end of Harry Potter" and also "because it helps me to remember that everything will be well at the end." She remembers these words not as *merely* four words but as the conclusion of the story; indeed, this quote itself is so threadbare that it depends upon some mental reference, however, hazy to the long and arduous struggle of Harry and his wizard comrades against Voldemort. Only with the knowledge that this quote comes at the close of this adventure can it be interpreted as being about how "everything will be well *at the end*."

Memorized quotes, including book quotes, play an important metacognitive function in the lives of quoters. Stored in the recesses of one's mind, they can leap into action to "correct" one's perspective or behavior or to allay one's anxieties. The quotations buried in quoters' memories may be connected to their use of Tumblr. In the case of the quote about Celaena's "scars," that this quote was in fact one of two quotes on Tumblr from this book that originally drew her to read the book in the first place. Another quoter, RQ100 (f, 20, Chat), remembered this quote from Haruki Murakami, which she originally found floating on Tumblr:

*But what seems like a reasonable distance to one person might feel too far to somebody else.*

It has become an important and enduring part of her psyche:

It somehow relates and speaks about my behavior and way of thinking. I think about it almost all the time—being an unsocial person who has problems bonding with someone.

Memorization may also be aided by the "deep attention" devoted to quotes in the process of posting them to Tumblr. RQ7 (f, 25, Chat) who asserted that she does not have a "good memory," nonetheless claimed the following:

If I were asked: Tell me a quote from *inserts a certain book* there's always a quote, above the rest, that would come to my head.

She suggested that "the whole process of reading, preparing the quotes, posting them and seeing them around helps" her to remember them. "It also helps," she added, "if you actually like the quote."

In such cases, Tumblr as external memory and *actual* memory are intertwined. However, over and above such concrete threads that connect the digital world of Tumblr with the mental worlds of its users, these testimonials about the power of memorization suggest the degree to

which quotation, for quoters, *is not* just about Tumblr or even about books. Arguably, this way of returning to words in one's own mind, not merely for the pleasure of recollection but as a tool for self-care and even survival, is the "deepest" way of experiencing quotations yet described.

### Discussion

Once again, the vocabulary of "deep" versus "hyper" attention obscures as much as it reveals. Certainly it is fair to say that there are ways that quotation is connected to ways of reading that are traditional and may be assumed to be deep. Quotations are lures. In their beauty and expressivity, they capture readers who are reeled back to the original book by their own desire. And even for some reader-quoters, part of the charm of the quotation is its ability to become just another morsel text in an endless feed, a momentary pleasure. Yet between these two poles is an array of tactical acts of literacy that cry out for a different theorization of attention. Quotation may for some distract from reading; for others it may actually make the experience of reading "deeper" by demanding intense focus on a few words.

In fact, the same reader may experience quotation in ways that are more and less deep. Quoting is connected to literacy practices such as reading and writing but also a literacy practice that is sometimes forgotten: quotes from books, or parts of them, may be housed in memory where they become useful "reminders" of how to act or feel.

While I have used the term "deep" to describe the degree to which one "pays attention" to something, this word might also be used to describe the intense emotional affinity quoters ideally experience when they encounter words in a book that scream out to be quoted or those that already have been and which circulate on the network. In general, the ideal way of experiencing a quotation is not with a detached consideration, however deep, but with powerful feeling. At last count, I must also agree with the quoters' own suspicion that this particular form of self-expression and self-care reproduces a gendered logic that casts females as expressive, males as laconic.

Quotation can thus be seen as part of a wider constellation of characteristically-female forms of expression on social media, and it serves as a helpful counterexample to certain stereotypes about the ways that young women themselves seek attention online. As Derek Conrad Murray (2015) has argued, popular discourses about young women often target the ubiquitous practice of posting photos of oneself ("selfies") on social media as a clear indication of this population's overwhelming narcissism (as well as a litany of other psychological problems and ethical failures). Murray himself argues against this negative reading of the selfie, suggesting instead that these images, as acts of self-representation, may allow young women to "envision themselves anew" (pp. 512). While there may be some underlying similarities between the two practices, quotation is in certain ways the "anti-selfie." Whereas the selfie indexes its creator uniquely (her face, her clothes), the quotation both does and does not index the person who creates it or shares it. Yes, the words may express some feeling that resonates with her, but they are not *her* words. And ultimately the feeling expressed is not even *her* feeling, at least not hers alone, since the point of quoting is often to experience one's emotions as widely shared, even universal, and thus to feel "validated." Quoting can even be altruistic, as when one puts a quotation onto the network so that it may be found by one who needs it. In these ways, quotations are not about *me*; they are about the *us* that can be found in the *me*.

## CHAPTER 5 ANALYZING THE LINGUISTIC CHARACTER OF BOOK QUOTES

Phase III focuses on the *language of quotations*. It answers the question: What linguistic features characterize quotations on Tumblr, especially those from print-based media, and how do these features relate to users' purposes for quoting?

Studying the linguistic aspects of quotations is another way of getting at the practice of quotation. It can reveal the literary values that users bring to texts in order to determine that certain words are worth more attention than others. After all, quotations are not like the works from which they came. *The Grapes of Wrath* and "To His Coy Mistress" each has one author; a quotation from either of these works, however, has two. Somebody, somewhere, at some point had to make choices (silent, but creative) about which words should be quoted. Quoting is fundamentally about *attention* in the sense that, by quoting some words and not others, quoters have implicitly directed their (and others') attention toward those particular words. Yet the relationship between quotation and attention is in fact more complex than this, since quotations themselves may be produced for different purposes and thus may expect or solicit different forms of attention. This chapter investigates how different styles of quotes seem to expect attention that is more or less "deep" and how certain kinds of quotations are more "quotable" than others. Furthermore, it reveals styles of quotation that do not directly fit into the "deep"/"hyper" binary, and it suggests that quotations are designed for quoters' expressive purposes (described in the previous chapter).

### Methodological Overture

#### Distant and Close Linguistic Analysis

The form of discourse analysis carried out in this phase often borrows methods from "corpus linguistics."<sup>42</sup> As its name implies, the first step in this kind of linguistic analysis is the creation of a *corpus* of real language data, written or spoken, or the acquisition of a pre-existing corpus, such as the Brown Corpus, a multi-genre corpus of written English that is commonly used in computational linguistic research. In such analysis, the corpus is often large enough to facilitate and require computational analysis. Corpus linguists have traditionally used techniques of quantification, such as counting the number of certain words or other discourse units within a corpus and testing hypotheses about the relative frequencies of words, though they may also use other techniques of computational assistance, such as automatic concordancing (Biber, Conrad, & Reppen, 1998; McEnery & Wilson, 2001). Computer-assisted research has long been a mainstream approach within Discourse Analysis, and the boundary between this field and

---

<sup>42</sup>While "corpus linguistics," as an academic field, is somewhat distinct from "discourse analysis," the line between them is a blurry one indeed, and has been since at least the early 1990s (a history traced by Baker and McEnery, 2015). *The Handbook of Discourse Analysis* includes a chapter in which Stubbs (2001) reviews a variety of corpus analytic techniques, including the identification of key words. Following this as well as the examples contained in Baker and McEnery (2016; see also Baker, 2006), I take it for granted that corpus linguistic methods, including those dealing with units as small as words, can be used as part of discourse analysis.

Corpus Linguistics is a porous one (Stubbs, 1996).

The methodological approach of Corpus Linguistics may be helpfully contrasted with a related field, Natural Language Processing (NLP). While some Corpus Linguistic research occurs under the banner of NLP, this term describes what is first a computer scientific endeavor to model language data with an eye toward specific tasks, such as machine translation or automatic summarization. One other difference is important: while NLP *generally* cleaves more to a scientific approach and privileges quantitative validation techniques (such as comparing the predictive performance of models of language), Corpus Linguistics is a more humanistic enterprise, one that is more comfortable with using computational analysis as a tool for producing new interpretations of a text. As Teubert (2005) argues in his manifesto of Corpus Linguistics as a humanistic discipline, "interpretation, not verification, is the proper response to the quest for meaning." In this sense, Corpus Linguistics has something in common with certain approaches to the "Digital Humanities" (e.g. Ramsay, 2011).

**Linguistics and "Computational Textual Ethnography."** While this Phase is inspired by computational linguistic research, especially as it relates to quotations (Bendersky & Smith, 2012; Danescu-Niculescu-Mizil, Cheng, Kleinberg, & Lee, 2012), it is also fundamentally different in its approach. This Phase uses linguistic analysis in the service of a broader ethnographic perspective. It does so most notably by placing into conversation with the findings of the earlier phases (especially Phase II), how the linguistic nature of quotations confirms or complicates quoter's own descriptions of the purposes of quotation. Also, unlike purely linguistic studies mentioned above, it does not tacitly assume some corpus of quotations as being representative of quotations in general. Rather it compares corpora of quotations to illuminate the ways that Tumblr quotations are different than other quotations. Linguistic differences between corpora of the same genre, quotations, may shed light on differences in the ways that specific communities use this genre.

My central assumption is that the formal qualities of quotations are linked to the matter of attention in two ways. First, to quote a passage from a book is a straightforward sign that it has caught the quoter's attention. In a more general sense, the formal and linguistic qualities of texts in general and quotes in particular provide clues about their expected audiences, even the kinds of attention that they expect or "solicit."

**Literary language.** In terms of object, discourse analysis has traditionally concerned itself with non-literary, even mundane textual artifacts, such as advertisements (Van Dijk, 1985), transcriptions of conversations (Gee, 2014), and bureaucratic texts (Fairclough, 1993). Quotations, especially those from literary texts, are not examples of everyday language but of "literary" language. The concept of literature poses a particular problem for discourse analysis: is literature just another set of discourses among many discourses? This question is echoed in the division of disciplines: discourse analysis (as a methodological perspective that is essentially a branch of linguistics) is separate from literary analysis. Still, there is a long tradition of considering what is or is not particular about literary texts *as* discourse—that is to say, those who approach literary texts linguistically (see Chatman, 1980; Jakobson, 1960; Todorov, 1981; Traugott, 1973; Pratt, 1977).

For my purposes, the key question is what sort of *vocabulary* is useful for describing fragments of literary objects. On the one hand, it is possible to use a vocabulary derived from linguistics, and discourse analysis more specifically, to describe literary artifacts. Halliday

(1971), for example, used his own categories of Systemic Functional Linguistics to analyze a novel by Golding. Widdowson (1992) analyzed the style of poetry in terms of familiar linguistic units that are by no means specific to poetry (e.g. noun phrases, clauses, etc.) Yet the same units of linguistics that can be used to describe everyday language are not necessarily sufficient for describing poetic language. Particularly relevant to this problem as it relates to quotations is the notion of *figurality*. Classical rhetoricians, and their Early Modern interpreters, described the structures and patterns that make language beautiful, or simply effective for a certain purpose (such as persuasion) (Burton, 2015). Renaissance school books listed anywhere from 65 to nearly 200 individual *figura*—figures—these ways of using speech that are, according to Quintilian, "artfully varied from common usage" (Corbett, 1965). Taking a more modern, linguistic perspective, Todorov (1981) defines a figure as any relation of two or more words that "permits itself to be described as such" (pp. 21). In a more general sense, Todorov's message is that a key part of the analysis of literary text as "discourse" is noticing and describing artful patterns of language, a task that depends upon the use of linguistic vocabulary but also extends beyond it.

Yet scholars such as Todorov and Jakobson, whose analyses are rooted in linguistics, are not unique in their ability to discuss what is particularly "poetic" about literary language. Literary critics and even writers themselves have also provided lucid, compelling, and precise accounts of what gives literary language its aesthetic power. Particularly relevant examples include Fussell's (1979) analysis of poetic meter and form and Voigt's (2009) account of poetic syntax. While these studies of poetic language attend to the particular organization of literary language, they approach these linguistic structures in a way that is more "humanistic" and linguistic. A good example is Fussell's discussion of the "volta" (or *turn*) of a sonnet:

If the two parts of the sonnet, although quantitatively unequal, can be said to resemble the two sides of an equation, then the turn is something like an equals sign: it sets into action the relationship between two things, and triggers a total statement. We may even suggest that one of the emotional archetypes of the Petrarchan sonnet structure is the pattern of sexual pressure and release. Surely no sonnet succeeds as a sonnet that does not execute at the turn something analogous to the general kinds of "release" with which the reader's muscles and nervous system are familiar. (pp. 116)

Fussell connects *form* to plausible *aesthetic effect*: the linguistic aspects of the sonnet produce and release tension. Likewise, Voigt's captivating account of the ways that syntax and enjambments work against each other in lyric poems weaves together a formalist account of the linguistic minutiae of verse with the effects the poem produces.

My own perspective on performing "discourse analysis" on literary texts is a middle position: I believe that the vocabulary of linguistics *proper* is useful for describing literary texts, but it is hardly sufficient. In my analysis of quotations qua literary texts I make use of pre-existing linguistic categories, yet I also at times name new *figura* a la Todorov, sometimes with reference to the vocabulary that critics have used to describe literary language in particular.

## **Data, Features, and Methods**

### **Data**

This chapter investigates several corpora of quotations from Tumblr, drawing comparisons between them and against "comparison" corpora consisting of texts from other

sources and networks.

### **Tumblr Data.**

*Book Quotes (BQ)/General Quotes (GQ) Corpora.* From the BQ Unfiltered Corpus, I created a corpus specifically designed for linguistic analysis. For each individual user in the dataset, I randomly selected one quotation. I used data-processing techniques to remove repeat quotations and to clean the data, including removing non-English quotations and extracting sources that users had included in the text of the quotation. (Users occasionally append a source such as " - John Green" to the body of a quote rather than typing it in the provided source field.) As much as possible this dataset includes quotations *from* books, not those *about* books. To remove quotations *about* books,<sup>43</sup> I did not include any quotations that contain certain words related to books (e.g. "book", "reading", etc.). After cleaning I was left with a corpus of 5,357 quotations, each one from a different quoter. I did the same for the GQ Unfiltered Corpus but limited the size of this corpus to 10,000 quotes, chosen at random.

*BQ 1-Sent/2-Sent Corpora.* To create a corpus for the analysis of common syntactic patterns, I created a sub-corpus consisting of all quotes from BQ Corpus consisting of exactly 1 sentence, as determined by NLTK's sentence tokenizer (n=2,690).<sup>44</sup> Building on my own work with Hearst (2016), I investigated discourse-level patterns in quotations (i.e., patterns above the level of the sentence). For this purpose, I created a corpus of all quotations in the BQ Corpus consisting of exactly two sentences (n=1,171).

*Pop/Unpop Corpus.* To investigate what makes certain quotes more popular than others, I created a paired corpus consisting of popular and unpopular quotations attributed to the same author. The names of authors were taken from the user-generated list of "Must-Read Classics" on the website Goodreads.<sup>45</sup> Each name was entered as a search query in the Tumblr API, returning all posts of the quote-type containing this name as a tag. For each author, quotes were sorted according to the number of times the quote was repeated by unique users. I defined "popular" quotes as those that had been repeated by at least three different users. "Unpopular" quotes were those that appeared only once. For each individual author, I made pairs of popular and unpopular quotes. I constructed the final dataset from these pairs, limited to 5 pairs maximum per author. This consisted of 1,284 individual quotations (642 Popular, 642 Unpopular) sourced from 153 authors.

*BQ Pop/Unpop Blogs Corpora.* To investigate differences between popular and unpopular blogs, I created a BQ Unfiltered Corpus to create two corpora consisting of posts from popular and unpopular blogs. Blogs whose posts were represented in the BQ Unfiltered Corpus were ranked according to the mean number of notes (likes and reblogs) their posts had received. I excluded any account with fewer than 10 quotations in the corpus because I wanted to

---

<sup>43</sup> For example, e.g. "'Never trust anyone who has not brought a book with them' (Lemony Snicket)

<sup>44</sup> The sentence tokenizer takes a string of characters (such as, "This is a sentence. And this is another.") and breaks it into sentences: (a) "This is a sentence."(b) "And this is another."

<sup>45</sup>[http://www.goodreads.com/list/show/449.Must\\_Read\\_Classics](http://www.goodreads.com/list/show/449.Must_Read_Classics)



investigate differences between the "practices" of blogs—i.e., what kinds of quotes popular versus unpopular blogs *tend* to post. This was an arbitrary cut-off but it seemed proper to ensure some sort of minimum. To use a metaphor from another field: it would be unwise to judge a batter as good or bad based on this player's performance in a single plate appearance. And certainly it would be unfair to compare a batter with one plate appearance to one with many.

The BQ Pop Blogs Corpus (n=1,000) consists of 10 quotes from each of the 100 most popular blogs according to this ranking. The average number of notes for these users' posts was 243.8. Likewise BQ Unpop Blogs Corpus (n=1,000) consists of 10 quotes from each of the 100 least popular blogs according to this ranking. The average number of notes for these users' posts was 1.3.

### **Comparison Corpora.**

*Brown Corpus.* Despite the fact that this corpus consists of texts from 1961, the Brown Corpus (Francis & Kucera, 1979) is a standard and often-used corpus for linguistic research today. Consisting of about a million individual tokens or "words," it represents written American English in a wide variety of genres, including both fiction and nonfiction. I used the version of the Brown Corpus that is packaged with the Python Natural Language Toolkit (Bird, Klein, & Loper, 2009).

*Brown Lit Corpus.* For my research, I created a subsample of the Brown corpus. With this corpus I attempted to create a corpus of generally "literary" (rather than expository or informational) texts, more in keeping with the types of texts from which Tumblr quotations are usually extracted. It consisted of the following categories: *belles\_lettres*, *fiction*, *mystery*, *science\_fiction*, *adventure*, *romance*, and *humor*. My goal in creating this corpus was to make a corpus that was, in terms of genre, more closely aligned with those Tumblr quotations from books, making it a fairer comparison.

*Brown Lit 1-Sent Corpus.* This corpus is the "twin" corpus of the BQ 1-Sent Corpus. From the Brown Lit Corpus I randomly selected paragraphs and from each paragraph randomly selected one sentence. I randomly selected from these to create a corpus the same size as the BQ 1-Sent corpus (n=2,690).

*Twitter Quotes Corpus.* Though this is a study of Tumblr quotes, as I conducted preliminary analysis and reached the findings explored in other chapters, I found it necessary to make some basic comparisons between Tumblr and another popular network that included quotations, in order to get a sense of whether my findings about Tumblr quotes could be findings about quotations in general. I selected Twitter because a respondent to my questionnaire (Chapter 4) suggested that Twitter (like Facebook) may be less "emotionally raw" than Tumblr. Practical reasons were also considered: Facebook is difficult to investigate, for instance, as data there is not generally public, while Twitter offers researchers an API. I also knew from previous experience on Twitter that quotes circulated on that site. Tweets tagged "#quote" were downloaded using the Twitter Search API. Twitter's API is much more restrictive than Tumblr's. For instance, it does not allow for searches over arbitrary periods of time. Searches are limited to the past seven days. In practice, I found the amount of data to be overwhelming, and so I sampled quotes from two days: 20,009 posts from January 23, 2017 and 20,786 posts from April

30, 2017. (Many of these were "retweets.")<sup>46</sup> After filtering out retweets and photo-type posts (containing a quote in an image rather than in text form) and cleaning the data to remove hashtags, links, and usernames, I chose one quote from each unique user. This left a corpus of 1,112 Twitter quotes.

## Features

The computational analysis in this Phase, as a form of "machine reading," took advantage of the ways that linguistic features in text can be formalized and thus quantified.

Certain features were designed for exploratory analysis of quotations and to not make *a priori* assumptions about what makes something quotable. Others, however, I investigated very specifically because they allow me to pursue specific assumptions about the purposes of quotations and the kinds of attention they solicit.

**Tokens.** The simplest features are words. In computational linguistic jargon, each countable instance of a word is called a "token." "Tokenization" is the process of breaking up a "string" (a series of characters) into a list of countable tokens. I used the NLTK tokenizer. In tokenization, punctuation marks are themselves tokens.

**Subject Nouns.** A dependency parser (Honnibal & Johnson, 2015) represents a sentence as a directed graph in which each word exists in one or more of a finite number of dependency relations to other words in the sentence. One word, labeled ROOT, is considered the core of the sentence. This is usually the central verb. In sentences with a nominal subject that is the agent of this main verb, this subject is connected to this central verb via the NSUBJ relation. For my purposes, I define a "subject noun" as the noun that takes the position in a sentence parsed in this way. This definition does not capture the entire subject of a sentence when this subject consists of a noun phrase longer than one word. To create this feature, I used the dependency parser offered in the SpaCy<sup>47</sup> suite for Python.

**High-level syntax patterns.** The Stanford Parser (Klein & Manning, 2003) parses a sentence according to context-free grammars, representing it as a tree structure. Nodes of the tree may be clauses, phrases, parts of speech, or individual words (called "terminals," since the tree "terminates" at these outer leaves). Each node may have at most one parent node. Following Feng, Banerjee, and Choi (2012), I created a feature by stringing together the nodes two levels below the ROOT level. For instance, the following sentences

*If youth is a defect, it is one that we outgrow too soon.*<sup>48</sup>

*If loving you is wrong, I don't want to be right.*

may both be represented as a *SBAR+NP+VP+.*, a subordinate clause followed by a comma, a noun phrase, a verb phrase, and finally a terminal punctuation. This feature offers an abstract representation of a sentence's overall grammatical structure, making it possible to recognize the latent similarities between sentences that have subtly different syntactic structures. As explained below, I only applied this feature to single-sentence quotations and non-quotations.

---

<sup>46</sup> Retweets are akin to Tumblr "reblogs." While Tumblr's API does not return "reblogs," Twitter's API returns "retweets."

<sup>47</sup> spacy.io

<sup>48</sup> Robert Lowell

**Length.** The length of a quotation (in sentences and in characters) is a simple feature, yet it directly gets at the matter of attention. I assume that a longer quotation quite literally demands more attention than a shorter one; in general it takes longer to read a quotation consisting of 1,000 characters than one consisting of 140.

**Person names.** During initial exploration of the data, I became interested in the presence (and absence) of character names (e.g. "Hermione" or "Gandalf"). Character names in quotations are of interest because to *fully* understand a quote mentioning, for instance, Hermione, one must have at least passing knowledge of *Harry Potter*. Put another way, I assume that containing person names may make quotations less easily consumed via "hyper" attention since they could demand some "deep" knowledge of the source text.

**Imperatives.** I designed this feature under the suspicion that quotations, as "equipment for living" (Burke, 1973), may offer explicit guidance regarding how one should live life. I assume that imperatives, by commanding the reader, tend to provide exactly this sort of explicit guidance—as does the example quotation from Auden: "If thou must choose / Between the chances, choose the odd." To find sentences in the imperative mood, I relied once again on the SpaCy dependency parser. In particular, I consider a statement imperative if it contains that is a (a) present-tense, (b) a dependent of ROOT, and (c) not in turn connected to any noun—in less-technical terms, a sentence without a main subject noun. This is a more robust approach than simply locating sentences that begin with a present tense verb, as imperatives may take more complex forms. In the quote from Auden, for instance, the present tense verb that is a dependent of ROOT would be "choose," coming after the subordinate clause. In the cases of quotations containing more than one sentence, I consider only whether the first sentence contains an imperative (making comparison between quotations of different lengths fairer).

**Repetition.** Intuitively, quotations are not just language but *beautiful* language. Analyzing beauty, of course, is just the sort of vague and deeply human (i.e. subjective) process that is difficult for computers. My approach, which offers only a shallow proxy of the ways that linguistic material can be organized for poetic effect, is to focus on identifying repetition within and between sentences. A small cornucopia of classical rhetorical features describe the different forms of linguistic repetition. Rather than try to capture any single feature of classical rhetoric (see Dubremetz & Nivre, 2015; Gawryjolek, 2009), I took a more general approach. I operationalized the concept of "repetition" simply: a quotation is considered to contain a repetition if any sequence of four tokens (a 4-gram) occurred within it at least twice. This is an admittedly strict definition: "*O what is that sound which so thrills the ear...O what is that light I see flashing so clear*" (Auden). Such repetitions were considered over the span of the entire quotation, regardless of its number of sentences.<sup>49</sup>

## Analyses

---

<sup>49</sup>Repetition of a single word may be used in ways that are obviously poetic but also in ways that are not. This also does not capture repetitions in which there are slight differences in a series of repeated words or more complex forms of repetition in which the order of words changes, such as *chiasmus*. My intuition in making this feature so strict was to favor precision over recall.

**Exploratory comparisons.** I used a technique developed by Rayson and Garside (2000) to explore linguistic data by surfacing "key" linguistic items, such as "key words." For their purposes, a key word is defined relatively. Given two corpora, A and B, a word is a "key word" in A if it appears much more frequently in A relative to B. To calculate this, Rayson and Garside suggest a log-likelihood-based comparison based on a contingency table representing the frequency of a word in A and B along with the total vocabulary of A and B. The word with the highest resulting log-likelihood value in A is its most "key" word. This word goes at the top of a list of all words in A, reranked according to this score. Rayson and Garside suggest a minimum log-likelihood value of 6.63 as statistically significant and therefore "key." As they argue, this feature is useful for comparing units in any two corpora, not just words. Below I describe the exploratory comparisons I made with this technique:

- **Key Words.** Using frequency-profiling, I compared the presence of tokens<sup>50</sup> and subject nouns. In the case of tokens, comparison was made between equal-sized corpora (141,560 tokens each) drawn from the BQ Corpus and the Brown Lit Corpus; the larger corpus was randomly sampled. I filtered out repeat tokens on a per-quote basis, so that a quote containing the token "golf" fifteen times would only contribute one instance of "golf." I excluded punctuation to focus key *words*.
- **Key Subjects.** In the case of subject nouns, comparison was made between the Brown Lit 1-Sent Corpus and the BQ 1-Sent Corpus.
- **Key High-Level Syntax.** Using frequency profiling, I compared the presence of high-level syntax features in the Brown Lit 1-Sent Corpus and the BQ 1-Sent Corpus.

To analyze which token occurred in either the first or second sentences of certain quotations in the BQ 2-Sent Corpus, I compared the times this token occurred in the first sentence of a quotation with the times it occurred in the second sentence using a chi-squared test.<sup>51</sup> I ranked the words according to chi-squared value, though I also list raw frequency and the number of times the more common the word is in either the first or second sentence (see Booten & Hearst, 2016). I also note the "disparity," which I define as the number of times more frequent the token is in its more frequent position (the first or second sentence). This method provides yet another way of understanding the latent structures of quotations, providing more clues as to what captures the attention of readers.

**Analyzing examples.** When using Rayson and Garside's method and the other method described immediately above, I focused on certain interesting candidate phenomena. For instance, imagine that the word with the highest log-likelihood score in corpus A is the word "king." This on its own tells me close to nothing, since words cannot be understood out of context. "King" might in this corpus refer to the monarch of some territory; if this is always the case, then we may rightly guess that the corpus is about British history or Arthurian legend. But the word "king" might just as well refer to a very large mattress, a chess piece, or a male termite. Most troubling of all: it may mean, in different examples, any of these definitions. In analyzing small samples drawn from the corpus (usually  $n=20$ ), my main concern was to judge the degree

---

<sup>50</sup> See Rowberry (2016) for a complementary exploration of "key words" in Kindle highlights.

<sup>51</sup> In this statistical comparison, the implicit null hypothesis is that a given word is equally likely to occur in the first sentence as the second sentence, much like coin that is assumed to land heads or tails with equal likelihood.

to which a linguistic feature captured in the corpus tended to be used.

In my close analysis of "key" features, I am able to make claims about the particular kinds of attention solicited by quotations that contain some of these features while also connecting to quoters' purposes for quoting.

**Hypothesis testing.** My underlying hypotheses were that brevity, use of repetitions, and use of imperatives contribute to "quotability," since they make a quotation suited for "hyper" attention, aesthetically pleasing, and useful as a "tool for living" (Burke, 1973). Likewise, I hypothesized that containing character names works against a quote's quotability by making it more difficult to consume by those who have not engaged "deeply" enough with the text to know about the characters (i.e. making it less easily "decontextualized"). I tested these more general hypotheses by exploring specific hypotheses about the linguistic characteristics of specific corpora:

- That, relative to the Brown Lit 1-Sent Corpus, the BQ 1-Sent Corpus would have more of certain features that I assumed to characterize quotability (imperative statements, repetition) and fewer of a feature that I assumed to make text less quotable (character names)
- That, relative the Unpop Quotes, Pop Quotes would have more repetition and imperative statements and fewer character names, and would also be shorter.
- That, relative to the BQ Unpop Blogs, BQ Pop Blogs would have more repetition and imperative statements and fewer character names, and would also be shorter.

I tested these hypotheses with the relevant statistical tools. For instance, in comparing the mean lengths of texts I used two-tailed t-test ( $\alpha=.05$ ). In testing whether a feature was more likely to occur in one corpus than the other I used a chi-squared test ( $\alpha=.05$ ). Once again, corpora were "re-equalled" before comparisons by randomly sampling from the larger corpus.

**Further Analyses.** Certain analyses were not planned from the beginning using *a priori* categories. Rather I undertook them in the flow of research in response to findings unearthed by exploratory methods in conversation with findings from the previous phase:

- After landing upon the importance of pronouns in quotations, I compared the ratios of different categories of personal pronouns (first, second, and third) in the BQ Corpus.
- While exploring the high-level syntax features that characterized the BQ Corpus, I found the need to further explore the data set by generalizing certain syntactic patterns yet further; using regular expressions, I developed several string-based patterns that captured similarities between various high-level syntax patterns and compared the frequency of matches in the BQ 1-Sent Corpus versus Brown Lit 1-Sent Corpus.
- Later in the analysis, I began to wonder whether certain features that seemed connected to specific purposes for quotations on Tumblr were indeed Tumblr-specific. To interrogate this, I compared quotations from Twitter with those from Tumblr. Specifically, I compared the Twitter Quotes Corpus to the 2,836 quotes in the BQ Corpus that were less than or equal to 130 characters and the 3,400 quotes chosen from a random sample of the very large GQ Corpus that were less than or equal to 130 characters. Though tweets can be up to 140 characters, many tweets contain links and hashtags (which I removed) making their effective length shorter. I compared the frequencies of certain kinds of pronouns, a comparison that emerged naturally from my results explained in this chapter. As a rough proxy of sentiment or affect, I also compared the number of sentences containing "positive"

or "negative" words. Here I used a pre-existing dictionary of 6,789 words labeled "positive" or "negative" (Hu & Liu, 2004). This basic analysis of sentiment in some sense probed the degree to which Tumblr may in fact be more "emotionally raw" than Twitter, as one respondent to the questionnaire suggested.

This Phase also relied on data gathered from the already-described questionnaire (from Phase I). One of this instrument's questions asked respondents to rank a quotation by Toni Morrison or Fyodor Dostoyevsky. Respondents were shown the quote either with authorial attribution or without it. This question helped explore the degree to which authorship matters in quoters' judgment of quotes. Users scored the quote's quotability on a Likert scale from 0 to 10, 10 being highest or most "quotable." The mean scores of the two versions were compared using a t-test. The goal of this question was to test whether or not an authorial brings a quotation a more favorable reception. In other words, I wanted to know whether quoters pay attention not just to words but also to the author of those words.

## Findings

### Length

A quotation has a common sense definition: it is a text that has been excerpted—*intentionally*—from some other text.<sup>52</sup> But a quotation is also a linguistic phenomenon, a genre. I begin by accounting for some of the most basic features of quotations as they occur on Tumblr, beginning with length.

Recall that quoters in Chapter 4 drew attention to the ways that a quotation should be short. But just how short? In my sample "book quote" or "book quotes" (BQ Corpus1) the average length of a quotation was 2.3 sentences and 160 characters long—a bit longer than a tweet. In addition, the most common length of quotation was one sentence; one sentence quotes were nearly twice as common as those consisting of two sentences (which, in turn, were nearly twice as common as those consisting of three sentences). For comparison, this is shorter than the average length of a paragraph in the Brown Corpus (3.9 sentences per paragraph)<sup>53</sup>

Although quotations were, as a rule, short, there were exceptions. The longest quotation in the sample unraveled over 38 sentences (an extended passage from Yan Martel's *Life of Pi*). Such extremely long quotations were, however, quite rare; 95% of quotations contained six sentences or fewer.

### Key Words & Key Subjects

Brevity is perhaps the feature of quotations that most unifies the texts in the corpus of book quotes, yet it is only the most basic. A quote might most often be one sentence, but not any sentence pulled randomly from a book would make a good quotation. What sorts of words tend to appear in quotations?

Table 1 presents the top 50 key words in the BQ Corpus. Table 2 presents the key subjects (nouns that are the main subject of the sentence according to dependency parsing).

---

<sup>52</sup>As I noted in Chapter 2 and again in Chapter 3, a Tumblr "quote" may sometimes violate this definition, as when users invent their own pithy "quotes," excerpted from nothing.

<sup>53</sup>Calculated for 7,500 randomly chosen paragraphs.

Table 1.

*Top 50 Key words in BQ Corpus relative to Brown Lit Corpus.*

<b>Rank</b>	<b>Token</b>	<b>Log-Likelihood</b>
1	love	489.42
2	is	471.53
3	you	430.40
4	your	290.65
5	people	278.24
6	are	251.19
7	life	225.85
8	it	206.85
9	things	190.71
10	like	175.30
11	my	169.91
12	someone	168.94
13	world	163.92
14	heart	161.82
15	're	154.18
16	me	146.37
17	can	143.83
18	that	143.02
19	i	142.52
20	to	139.47
21	sometimes	125.85
22	just	123.51
23	everything	121.76
24	do	109.78
25	feel	109.76
26	always	107.22
27	live	104.80
28	thing	102.32
29	want	97.67
30	the	97.65
31	will	94.67
32	know	94.40
33	be	94.39
34	never	89.48
35	yourself	88.63
36	because	88.52
37	n't	85.37
38	we	85.12
39	ca	84.14
40	if	83.66
41	person	78.98
42	does	77.44

43	something	77.12
44	all	76.17
45	how	71.43
46	really	71.06
47	what	65.64
48	ever	65.12
49	pain	64.64
50	think	64.51

Table 2.

*Key subjects in BQ 1-Sent Corpus relative to Brown Lit 1-Sent Corpus.*

<b>Rank</b>	<b>Token</b>	<b>Log-Likelihood</b>
1	you	84.83
2	i	79.07
3	people	45.36
4	love	28.88
5	we	22.76
6	life	22.18
7	it	19.82
8	world	16.64
9	words	13.86
10	everything	13.45
11	things	12.40
12	thing	10.89
13	nothing	9.70
14	god	8.32
15	beauty	6.93

**Big words for big questions.** Common to both Table 1 and Table 2 are a series of nouns: "people," "life," and "world."<sup>54</sup> In a broad sense, they might all be considered to be "abstract" and indeed *vague* nouns: time (not an hour), people (not any particular person, such as George Gershwin), life (the whole of it, not any particular part of it), and world (the whole of it, not any particular location). One of these, "life," is among the most popular hashtags for quotes in general as well as book quotes in particular (see Chapter 3). When these words occur as *subjects*, they embody the tendency of quotations to make grand statements about these abstractions. By examining a sample (n=20), I confirmed that these quotations tended to describe "people" in the sense of "persons unspecified as regards number, class, or identity."<sup>55</sup> In fifteen of these twenty, the subject "people" took a present tense verb:

*People always clap for the wrong things.* (Attrib. J.D. Salinger)

*Well, people aren't perfect, and we wouldn't like them if they were.* (Attrib. Sian James)

The same general trend was true for quotations in which "world" is the first subject: Despite its pockets of darkness, the **world** can be a beautiful place.

<sup>54</sup> "Love" may also be considered an "abstract" term; I discuss it below in terms of the fact that it specifies an emotion, which I suggest is a just as salient aspect of this term.

<sup>55</sup>The first definition given by the *Oxford English Dictionary*.



*The world is not a wish-granting factory.* (Attrib. John Green)

And "life":

*Life is real and sometimes it's ugly and you just have to learn how to cope.* (Attrib. Colleen Hoover)

*Life isn't a support system for art.* (Attrib. Stephen King)

These quotations could be considered *adages* or *sayings*. They present some general, enduring truth about a vast, important concept.<sup>56</sup> Yet this is far from the only form that quotations can take. In this respect, it is instructive to compare those quotations in which a word occurs as the first subject and those in which those words simply appear.

**Everything or nothing.** Quotations do not shy away from making big claims about big concepts. A similar rhetorical tendency may be seen in the relative popularity in book quotes of the words "everything," "nothing," and "all." These words operate at the extremes—yet an opposite trend may also be seen in the fact that "sometimes" is also a key word.

*Nothing in life is ever as simple as we imagine.* (Attrib. Arthur Golden)

*Everyone dies alone, Eragon.* (Attrib. Christopher Paolini)

These are quotations that tell a "truth" about life. They are "generalizable" in that they apply to unlimited subjects and situations. "Nothing in life" can be a relationship, an athletic pursuit, or a work endeavor. "All of us" and "everyone" includes, of course, every possible reader.<sup>57</sup>

Such quotations also perhaps indicate the way in which quotations, or at least some of them, tend to create "wisdom" (declared truths about reality) that avoid ambiguities. Certain deaths, one might argue, are lonelier than others. And ethicists, priests, and judges are all in the business of determining exactly who is "imperfect" and to what degree. The fact that words like "everything" and "nothing" are characteristic of quotations suggests that—or at least some of them—deal in not just generalizations but absolutes.

At the same time, it is possible to see the opposing trend in the fact that "sometimes"—not "always"—is also a key word. Yet "sometimes" also lends itself to generalization. Quotations containing this word describe not a consistent state but a predictably recurring one:

*Maybe sometimes we can only see the truth about ourselves if someone shows us where to look.* (Attrib. Michelle Hodkin)

*Sometimes, those who put up the biggest shields are those who are protecting the biggest hearts.* (Attrib. Penelope Ward)

Either of these quotations would be sufficient and meaningful without the word "sometimes." The inclusion of this word *hedges*; the truth the quote offers is not necessarily applicable to every situation. The hedge is indeed emphasized by the first quote's "Maybe," as if the message it declares is not guaranteed to be truthful at all. It is up to the reader, applying the quote to her own life, to determine if this moment is one of these "sometimes." Perhaps these hedges also contribute to a tone of uncertainty, as if the speaker of the quote is in the process of grasping a hard-won epiphany about life.

**Copular "is".** The word "is" is characteristic of book quotes (Table 1). Clearly the word

---

<sup>56</sup> See Bendersky & Smith (2012) and Rowberry (2016) for different analysis of abstractions in quotations.

<sup>57</sup> These terms echo the "quantifiers" that Bendersky & Smith, 2012, found to be predictive of quotability.

"is" lends itself to all manner of sentences expressing all sorts of sentiments. Still, in a sample of 20 quotations containing "is", 14 were used in a copular manner, linking the subject of a clause with a description of that subject. These sentences might be called "equation," since they take the implicit form X=Y:

*After all, to a well-organized mind, death is but the next great adventure.*

*The purpose of a storyteller is not to tell you how to think, but to give you questions to think upon.* (Attrib. J.K. Rowling)

Sometimes this "equation" was buried in the thoughts of a character:

*I learned that to love is to destroy, and that to be loved is to be destroyed.* (Attrib. Brandon Sanderson)

But the fact remains that the equation itself is in the present tense. The fact that "is" is a key word in book quotes is indicative of the ways that these quotations do indeed sometimes function as adages that offer timeless, transcendent definitions.

**Emotions.** In Chapter 3 it was found that many of the top hashtags that quoters apply to quotes—quotes in general as well as book quotes—are related to emotions, including negative ones such as "#sadness." Table 1 also draws attention to certain words that in various ways seem "emotional." These include "love," "heart," "pain," and "happiness." Inspection of samples of the ways that these terms were used in the corpus confirms that quotations often impart some truth about a particular emotion.

In the case of a random sample of twenty quotations in which "love" is the subject of the first sentence of the quotation, for instance, all of these referred to love as an abstract concept, though sometimes with qualifiers ("true love" or "a man's love").

*Love is a secret society, a community for two.* (Attrib. Melvin Burgess)

*True love is the greatest leap of faith there is.* (Attrib. Mia Sheridan)

Only five out of these twenty occurred in the past tense. All but one of the rest (in the future tense) were in the present tense. They tended to describe what love *is*—its enduring, transcendent, trans-historical essence.

In the case of the word "pain," half of a sample of twenty instances of quotes containing this word from the BQ Corpus explicitly or implicitly used this word to mean emotional (rather than physical) pain:

Sleep relieved me from the **pain** of reflection. (Attrib. Mary Shelley)

The remaining instances of "pain" were various and did not fit into a single category. Several were vague, though perhaps were more likely to refer to emotional rather than physical pain.<sup>58</sup>

The word "heart," in all but one of twenty sample quotes, meant not a physical organ but rather one's capacity to feel emotions, sometimes especially romantic or desirous ones. Yet this is a fragile and sometimes troublesome capacity:

*The heart can function with the wound, but it never entirely heals.* (Attrib. Janet Gurtler)

*The heart doesn't ask permission. It is singularly unconcerned with the qualifications of those it chooses to love. It mocks the intellect, it subjugates reason, and it holds hostage the will to survive.* (Attrib. Connie Brockway)

*The heart makes its choices without weighing the consequences. It doesn't look ahead to the lonely nights that follow.* (Attrib. Tess Gerritsen)

---

<sup>58</sup>For instance, "She was in pain and I loved her, sort of loved her, I guess, so I kind of had to love her pain, too."

Emotions figure strongly in quotations. Sometimes, as the examples above suggest, they take the form of quotations *about* love, *about* pain. But words like "pain" and "heart" can also enter into quotations without being the main subject of those quotes:

*If I survived the **pain** of losing him, I guess I can handle anything.* (Attrib. Josephine Angelini)

*In her eyes, I see my **heart**. In her breath, I hear my soul. She is my land. She is my kin. My love.* (Attrib. Pierce Brown)

These quotations are not *about* emotions so much as they are narrative expressions of those emotions in the context of quotations that invite the reader to identify with the "I," a feature explored in greater detail next.

**Pronouns.** "You" is a key word and also a key subject of quotations. In a sample of instances of book quotes (n=20) containing you, "you" was most often a *general* you (14 out of a total 20). Such quotations could take the form of an exhortation, even if it is not in the imperative mood:

***You** want to be good for the ones **you** love, because **you** know that your time with them will end up being too short, no matter how long it is.* (Attrib. Stephen King)

*She realized that **you** can't be weak when it comes to killing: **you** have to be strong or it just causes more sorrow.* (Attrib. Laura Esquivel)

***You** have to jump off cliffs and build your wings on the way down.* (Attrib. Ray Bradbury)

Such quotations could be reworded to remove the "you" with little change in meaning: "It is important to be good for the ones you love...", "One has to be strong...", and "One must jump off cliffs...". This kind of "you," I suggest, is a not very surprising characteristic. They mark quotations' tendency to proffer wisdom, implicitly invoking the reader as (part of) the general *you*.

The remaining quotes, however, referred to a not general—but also not specific—*you*. And 4 referred to a "you" that existed in relationship to an unnamed or implicit first person, often an "I." For instance:

***You** and I are nothing but wild beasts wearing human skins.* (Attrib. Sarah J. Maas)

*And if **you** can't say yes, answer anyway. Because I'd rather live with the answer, than die with the question.* (Attrib. Ian S. Thomas)

Yet, though the "you" and "I" imply a specific relationship between two people, these pronouns are still vague or "general" in the sense that the "you" and "I" are left open, unspecified. It is scarcely possible to glean from these quotations anything about these relationships, not even the gender of the participants.

This second kind of quotation—let us call it the "vague pronoun" quote—is in fact more characteristic of all quotations containing "I," especially those in which "I" is the subject. In analyzing a sample of quotations in which "I" is the subject, I drew a distinction between those in which the "I" seemed to refer to a specific person and those that did not—those in which the "I" is left vague. For instance, this quotation seems to be an example of the first:

*I swear, Six is going to kill us, or worse, maybe she's about to be killed by a swarm of Mogs and we're here lying in the grass about to go through a scene from *Romeo and Juliet*.* (Attrib. Pittacus Lore)

Compare that quotation to these:

*I would rather be numb than stand here in the light of a sun that can never chase the chill away.* (Attrib. Annie Kaufman & Jay Kristoff)

*I can now see everything falling to pieces before my eyes.* (Attrib. Ian Thomas Curtis)  
*I kissed him and let my lips rest against his so that our breath mingled and the tears from my eyes became salt on his skin, and I told myself that, somewhere, tiny particles of him would become tiny particles of me, ingested, swallowed, alive, perpetual.* (Attrib. Jojo Moyes)

*If I could avert my events from all the kissing people ever, I'm positive that my life would be at least 2 percent more fulfilling.* (Attrib. Julie Murphy)

Such quotations<sup>59</sup> do not present a general truth but rather a speaker's particular view on what is often a trying or painful reality. On a linguistic level, they are characterized by a particularly interesting use of *deixis*. "Deictic" words include pronouns, such as "I," but also words such as "here"; these words must be understood in reference to some other context, or they do not make sense (Segal & others, 1995). In the context of stories, such words make sense in reference to other elements of the fictional world. (And so that quotation about the "Mogs" and someone named "Six" makes full sense only if the reader understands this part of the story.) The "vague pronoun" quotes, extracted from their source text, do not supply the necessary context to meaningfully understand "I" and "you" in terms of the storyworld. At the same time, they do not include any other narrative that requires knowledge of the source text's narrative. Take the quotation that expresses a desire to be "numb"; certainly a "desire to be numb" is not a mysterious sentiment even in this stand-alone utterance. I suggest that the deictic pronouns here are left "open," demanding that the reader ascribe some meaning to the them.

Recall that, in data presented in Chapter 4, quoter often suggested that a good quotation is one that somehow expresses a sentiment that they themselves do not know how to put into words. They are tools of *self-expression*. Certain kind of pronouns, one quoter suggested in a follow-up interview, can make certain work more quotable:

I think it's no mistake on John Green's part that his writing tends to be either in first person or his most quoted lines within dialogue. Younger people who read it can envision themselves speaking like that.

The present analysis suggests that this is a more general phenomenon of quotations. They invite the quoter to identify with the "I," and sometimes to project some other person's identity onto the "you," filling the void left by these deictic terms. (Perhaps the "you" becomes a former friend or lover.) It is not surprising, especially in light of the general importance of affect (including negative emotions) in quotations, that these quotations are sometimes used to voice anxiety, misery, or heartache. Such quotations are a bit like mini-diaries in that they preserve confidential emotions. Yet while diaries maintain their secrets by virtue of being private (hidden somewhere safe, or protected by lock and key), these quotations express a quoter's private emotions in public, yet without any real reference to the quoter's specific life. As another quoter from Chapter 4 explained there is "a built-in barrier" between the quoter and the quote; it can express the quoter's emotions while also preserving a plausible deniability: "you can say, 'It's just a book.'" In these "vague pronoun" quotes, the "I" both is and is not the speaker; such quotations are like masks, except that they reveal even as they occlude.

*Gendered vs. non-gendered personal pronouns.* The quoter referenced above drew attention to the importance of "first person" pronouns. Furthermore, recall that in Chapter 4

---

<sup>59</sup>Though there were some ambiguous cases, these were the most common in my sample of 20; a conservative classification judged 15 of the quotations to be of this nature.

another quoter mentioned that a good quotation is one in which "gender-specific pronouns are neutral or not present." This characterization makes sense in light of the concept of *deixis*; a quotation about an "I" who romantically loves a "she" would not be as immediately useful to those readers who instead would feel romantic love for a "he." As a further investigation, I analyzed the BQ Corpus to determine the relative frequencies of quotations containing first person pronouns (such as "I" and "ours"), second person pronouns ("you" and "yours"), and gendered pronouns (such as "he" and "hers"). The raw frequencies of different types of pronouns in the BQ Corpus combinations suggests that non-gendered quotations are also the most common type (Table 3). In this table, "Only First" means those containing first person pronouns but not second person pronouns or gendered pronouns. The same pattern applies to the following two categories, "Only Second" and "Only Gendered."

Table 3.

*Percentage of quotes containing pronouns in the BQ Corpus.*

<b>Pronouns</b>	<b>Percentage Containing</b>
First	49%
Second	34%
Gendered	24%
Only First	26%
Only Second	17%
Only Gendered	10%

These measurements also confirm that non-gendered personal pronouns (first person pronouns but also second person pronouns) are generally more frequently used than specifically gendered pronouns. Assuming that non-gendered pronouns make it easier to "identify with" a quotation than gendered pronouns, this finding suggests that quotations are more frequently designed to be "identifiable" to readers. Still, it would be an exaggeration to say that quotations marked by gender are *forbidden*, even if they are somewhat less common—and, as I have argued, less "optimized" for the specific deictic function I have described.

### **Key Syntax**

There are interesting syntactic differences between quotations and non-quotations.

Table 4 depicts those high-level features which do have terminal punctuation (e.g. a period or question mark at the end of a sentence). Because they have this punctuation, they are "complete sentences." As I discovered in the case of high-level syntax patterns, some of those that proved to be relatively common in the BQ 1-Sent Corpus compared to the Brown Lit 1-Sent Corpus were due to the messy nature of the user-typed Tumblr data relative to the Brown Corpus data. For instance, some quotations from Tumblr do not include terminal punctuation. In my presentation of this data, I omit those high-level features that seem to be due to this messiness (i.e. high-level features that contain no terminal punctuation).<sup>60</sup> I focus my analysis on the first three patterns. After briefly analyzing all three of them, I will point to an underlying similarity between them—what I am calling the "mini-volta."

---

<sup>60</sup>Hence BQ 1-Sent Corpus has many instances of a high-level syntax feature such as *NP + VP*, a noun-phrase followed by a verb-phrase but missing terminal punctuation.

Table 4.

*Key high-level syntax patterns in the BQ I-Sent Corpus versus the Brown I-Sent Corpus.*

<b>Rank</b>	<b>Pattern</b>	<b>Log-Likelihood</b>
2	S + : + S + .	77.85
6	SBAR + , + ADVP + NP + VP + .	13.86
7	S + : + CC + S + .	13.45
12	SBAR + : + NP + VP + .	8.32
13	S + : + S + : + S + .	8.32
14	SBAR + , + WHADVP + SQ + .	8.32
15	S + : + S + : + CC + S + .	8.32
17	SBAR + VP + .	7.80
19	NP + VP + .	6.93

The first pattern, *S + : + S + .*, represents an independent clause followed by one of any of a certain number of non-terminal punctuation marks, such as a colon or em-dash, followed by another clause, followed by terminal punctuation. A few examples:

*Sometimes I think high school is one long hazing activity: if you are tough enough to survive this , they 'll let you become an adult.* (Attrib. Laurie Halse Anderson)

*No guys in his right mind would ever choose me when there are people like Hana in the world: It would be like settling for a stale cookie when what you really want is a big bowl of ice cream, whipped cream and cherries and chocolate sprinkles included .*

*That was what made fighting so easy - you could always choose death rather than captivity.* (Attrib. Cornelia Funke)

In terms of lower-order syntax, vocabulary, and length, these sentences are very different. What unites them is that they consist of two propositions that "pivot" in the middle. They are each really *two* sentences, joined together into one sentence through the conventions of punctuation. In each, there is a clear need for "elaboration" (Mann & Thompson, 1988) established by the first sentence. Why is high school like hazing? Why would no guy choose "me"? What made fighting so easy? Each of these creates an implicit lack of information that the second part of the quote satisfies. Quotations with the high-level pattern *S + : + CC + S + .* are similar, except with a coordinating conjunction after the punctuation that "bridges" the two main clauses:

*Regardless of sex, everyone loses something in a war...but the first casualty is always the TRUTH.* (Attrib. Brian K. Vaughn & Fiona Staples)

The pattern *SBAR + , + ADVP + NP + VP + .* represents a different yet also related discursive pattern in quotations: a subordinate clause, a comma, an adverbial phrase, a noun phrase, a verb phrase, and terminal punctuation. Below I place several examples:

*If you remember me, then I don't care if everyone else forgets.* (Attrib. Haruki Murakami)

*If I am to truly become an autonomous woman, then I must take over that role of being my own guardian.* (Attrib. Elizabeth Gilbert)

*If I never experienced how he could make me feel, then I wouldn't miss it after he's gone.* (Attrib. Colleen Hoover)

These are all conditional statements, in which the second clause fills in what will our would be true under the condition described in the first.

I generalized this pattern by finding all sentences beginning with "If" or "When" and also

containing a single comma.<sup>61</sup> A chi-squared test revealed this pattern too to be more common in the BQ 1-Sent Corpus than in the Brown Lit 1-Sent Corpus. These included those of a different high-level syntax pattern but with essentially the same underlying discursive pattern:

*When you forgive, you love.* (Attrib. Jon Krakauer)

*When life offers you a dream so far beyond any of your expectations, it's not reasonable to grieve when it ends.* (Attrib. Stephanie Meyer)

*If no one in the entire world cared about you, did you really exist at all?* (Attrib. Cassandra Clare)

Certainly there are finer distinctions that can be made between such quotations—between the uses of *if* and *when*, between declarative statements and questions. Some sound more like proverbs, while others seem more personal—plaintive questions or statements of heartbroken resignation. However, these quotations in general represent the pattern *If/When X, Y*. The *S + : + S + .* pattern, the *SBAR + , + ADVP + NP + VP + .* pattern (and the more general *If/When X, Y* pattern), and the *S + : + CC + S + .* pattern are in distinct, yet they share an underlying similarity. In general, they can each be thought of as consisting of *two parts*. In the case of the *S + : + S + .* pattern and the *S + : + CC + S + .* pattern, these two parts are independent clauses bridged by a colon or semicolon. In the case of the *If/When X, Y* pattern, the two parts are a dependent clause and an independent clause, also joined by punctuation. These quotations "pivot." They all have what might be called a "binary structure"—*X* followed by *Y*.

In that respect, they are a bit like certain types of jokes (i.e. "Why did the chicken cross the road?"..."To get to the other side."). But a different analogy might be made to the realm of poetry; in a sonnet, the "volta" (Italian for "turn") is the part where the sonnet's logical argument pivots, when the speaker complicates the ideas presented in the previous lines. As critics have observed, such a logical turn is crucial to the success of the sonnet (and lyric poems in general [Fussell, 1979]); a sonnet without a turn would be like a mystery story without a twist at the end. While we might think of lyric poetry as mere "beautiful language," the concept of the volta draws attention to the important role of *logic* in verse. In a sonnet, however, the volta occurs only after the bulk of the sonnet has already run its course—after eight lines in the Italian sonnet, twelve lines in the English. As the examples above demonstrate, what we might call a "mini-volta" may be generated much more expeditiously. Each of these quotes begins by positing some statement about reality in the first clause; this is just as quickly qualified, contradicted, complicated, or upended by the second. The quote "turns" on the central coordinating conjunction or on a punctuation that divides the quotation into its logical constituents.

**Other discursive patterns.** Table 5 presents features that tended to occur either in the 1st or 2nd sentence of the BQ 2-Sent Corpora in a statistically significant way according to a chi-squared test ( $p < .05$ ), ranked by the ratio according to chi-squared value comparing the frequency in S1 and S2. The bold number is the sentence with the higher count, S1 or S2.

---

<sup>61</sup>I found such sentences with the following regular expression:  $^(?:If|When)[^,]+,[^,]+[.?!]$\$ .

Table 5.

*Tokens that tend to appear either in the first or second sentence of a quotation (BQ 2-Sent Corpus).*

Rank	Token	S1	S2	Disparity	Chi-Squared Value
1	them	27	<b>78</b>	2.89	24.77
2	said	<b>48</b>	14	3.43	18.65
3	it	188	<b>278</b>	1.48	17.38
4	It	59	<b>111</b>	1.88	15.91
5	And	34	<b>74</b>	2.18	14.81
6	They	7	<b>29</b>	4.14	13.44
7	"	<b>59</b>	26	2.27	12.81
8	But	28	<b>61</b>	2.18	12.24
9	if	45	<b>84</b>	1.87	11.79
10	again	7	<b>27</b>	3.86	11.76
11	just	43	<b>80</b>	1.86	11.13
12	There	<b>39</b>	15	2.60	10.67
13	as	58	<b>96</b>	1.66	9.38
14	be	99	<b>146</b>	1.47	9.02
15	forever	5	<b>20</b>	4.00	9.00
16	have	74	<b>112</b>	1.51	7.76
17	made	7	<b>22</b>	3.14	7.76

The first trend that can be observed from these data is not surprising. Many of the words that occur in the second sentence of a quotation in some way signal a cohesive relationship (Halliday & Hasan, 2014) with the previous sentence. "They," "it"/"It," and "them" are pronouns that in many instances refer to a named subject in the previous sentence. Likewise, sentences that begin with "And" and "But" refer back to contradict or add to the previous sentence (perhaps at times performing another kind of "mini-volta").

Other tendencies, however, are more puzzling. Examining them can illuminate some of the underlying rhetorical tactics of quotations.

"*Just.*" Take, for instance, the rather ordinary word "just." Why would it appear in the second sentence more frequently than in the first? A few examples shed light on the typical rhetorical function of this humble little word:

*Sometimes we don't get second chances. Sometimes things **just** end.* (Attrib. Colleen Hoover)

*This is my heart. You can't **just** come here, and take it.* (Attrib. Ian Thomas)

*This was what true fear was-that you could never know other people, not completely.*

*That you were always **just** guessing blind.* (Attrib. Lauren Oliver)

Each of the above sentences would express the same facts about reality even if the word "just" were removed. Yet this does not mean that the word is of no consequence. Aarts (1996) suggests



that the word presents whatever claims about reality the sentence puts forward as "straightforward" and thus unobjectionable; the goal of the word is thus to convince the listener. More generally, this word seems to be "emphasize" a point, perhaps (as in the second quote) in a way that seems forcefully emotion. The present analysis suggests that this emphasis is best delayed until the *end* of the quotation, a sort of rhetorical punctuation.

"*There.*" "There," another seemingly milquetoast word, also plays a special rhetorical function in quotations.

*There's only one choice and the brave make it. Fly from the path.* (Attrib. Mark Z. Danielewski)

*There is power in beauty. That's the tragedy of it.* (Attrib. Amanda Filipacchi)

*There's no going back. And, when you can't go back, you have to worry only about the best way of moving forward.* (Attrib. Paulo Coelho)

These "There is" statements in quotations tend to make a confident and often broad claim about reality in general, as in "There is power in beauty.", or about what sounds like a specific but extremely vague situation ("There's no going back." and "There's only one choice..."). Sentences that use this construction to point to a *specific* reality (e.g. "There's a snake in my boot.") are relatively rare. In this sense, the "There" construction fits into a general pattern of vagueness and generalizability.

### Testing specific hypotheses

**BQ 1-Sent Corpus vs. Brown Lit 1-Sent Corpus.** I hypothesized that, relative to non-quotes, quotations would contain certain features that I thought may make quotations more quotable: imperative statements and repetition. I also hypothesized that they would contain fewer character names, a feature I thought would make quotations less generally quotable.

Comparing the BQ 1-Sent Corpus and the Brown Lit 1-Sent Corpus confirmed my hypotheses. The BQ 1-Sent Corpus contained approximately twice as many imperatives ( $\chi^2=5.15$ ,  $p<.05$ ) and eight times as many instances of repetition ( $\chi^2=60.8$ ,  $p<.001$ ). By contrast, the Brown Lit 1-Sent Corpus contained over five and a half times as many character names ( $\chi^2=43.7$ ,  $p<.001$ ).

**Pop Quotes Corpus vs. Unpop Quotes Corpus.** Similar results were found when comparing the Pop Quotes and Unpop Quotes corpora. According to chi-squared tests, Pop Quotes Corpus had nearly twice as many imperatives ( $\chi^2=4.14$ ,  $p<.05$ ), while the Unpop Quotes Corpus contained nearly twelve times as many character names ( $\chi^2=48.6$ ,  $p<.001$ ).

Quotes in the Pop Quotes Corpus tended to be considerably shorter than those in the Unpop Quotes Corpus. The average length of a Pop Quote was 155 characters, while the average length of an Unpop Quote was 280 characters. The difference in means was statistically significant according to a two-tailed t-test ( $t=8.43$ ,  $p<.001$ ).

Analysis did not support my hypothesis that popular quotes would have more repetition; in fact unpopular ones tended to have *more* repetition, though I suspect that this is largely due to the fact that unpopular quotations tended to be much longer (i.e. they contained more sequences of words that could possibly be repeated).

*A caveat: whose "good" quote?* Comparing quotes that are repeated with those that are

not implicitly places these texts on a spectrum between good and bad, between successful and unsuccessful.

But is this assumption not itself an uncritical repetition of the cultural logics of the media themselves? Tumblr posts (like posts on other networks) have built in metrics of popularity which are factored into the network's algorithm and which are presented to the user as visual proof of whether or not an item has succeeded or not. Never before have our utterances been so easily measured, and there is no doubt a temptation to regard a blog with few followers and a post with few likes as failures.

I linger here to apply some pressure to this assumption.

First, is an infrequently repeated quotation a "bad" quote? Consider the following quotation attributed to Goethe:

*Distance, my friend, is like futurity. A dim vastness is spread before our souls: the perceptions of our mind are as obscure as those of our vision; and we desire earnestly to surrender up our whole being, that it may be filled with the complete and perfect bliss of one glorious emotion. But alas! when we have attained our object, when the distant there becomes the present here, all is changed: we are as poor and circumscribed as ever, and our souls still languish for unattainable happiness.*

And another attributed to the same author:

*Knowing is not enough; we must apply. Willing is not enough; we must do.*

The most striking difference between these two passages (observable even before reading the first word of either) is first and foremost a difference in length, which, cast in a print medium, is a difference in shape: a veritable paragraph, a chunk, versus a nimble passage, a quick dagger of text.

The next striking difference is formal. The second quotation attributed to Goethe arrays itself in the stately garb of classical rhetoric. Its overall structure is one of *isocolon*, the use of a balanced parallelism. Each half of the equation fills in the blanks in this pattern:

X is not Y ; we must Z.

Within this pattern (complete with its own "mini-volta") is classical *antithesis*, in which a negative statement ("X is not Y") is juxtaposed with an often similarly structured positive statement ("we must Z"). It is as balanced as a classical *frieze*, but another word would throw off this linguistic equilibrium.

How can we describe the style of the longer quotation? To say that it is "prosaic" would be true but too vague. If the shorter quote contains itself within a rigorously efficient structure, the longer one tends towards the baroque details, especially in its use of adjectives: "dim vastness," "whole being," "perfect bliss," "glorious emotion," "unattainable happiness." It is not without repetition ("when we have attained..." and "when the distant..."), but not with matching restraint.

This longer quote is perhaps somewhat less easily "decontextualized" than the shorter one. (It is unclear from the passage that the "friend" addressed is Wilhelm, Werther's epistolary interlocutor.) Yet in other ways this longer quote is very much of the same type as the shorter one. Both deal almost exclusively in abstractions—"distance" and "futurity," "knowing" and "willing." The longer quote, while addressed to "dear friend," is like the shorter in that its main pronoun is a universalizing "we." Reading either of these quotes, the reader is recruited as part of this we. In the case of the first quote, it does not matter whether we the readers know that the words are from *The Sorrows of Young Werther*, that the words are those of Werther himself, that his friend's name is Wilhelm, that Werther makes this philosophical reflection while

simultaneously vouchsafing his love for the nearby Charlotte, or that this is a moment of contentment in a life that will end in suicide. In this case, the quote is what remains when such details are burned off.

These sorts of quotations that make abstract, generalizable statements about reality are perhaps the *main* type of quotations, yet quotations—especially those from books—can also fulfill another purpose. It is hard to find any universal wisdom in this quote attributed to Michael Ende's *Neverending Story*:

*This book was the Neverending Story, which he himself had read in the attic. Maybe his present adventures and sufferings were in the book even now. And maybe someone else would read the book someday - maybe someone was reading it at that very moment.*

Or in this one, from Kerouac:

*I had not met Japhy Ryder yet, I was about to the next week, or heard anything about "The Dharma Bums" although at this time I was a perfect Dharma Bum myself and considered myself a religious wanderer.*

Or in this one, from Tolkien:

*If the dwarves asked [Bilbo] what he was doing he answered: 'You said sitting on the doorstep and thinking would be my job, not to mention getting inside, so I am sitting and thinking.' But I am afraid he was not thinking much of the job, but of what lay beyond the blue distance, the quiet Western Land and the Hill and his hobbit-hole under it.*

Missing from these quotes are the linguistic markers of generalizability that were so striking in the two quotations from Goethe—the abstract nouns, and the universalizing "we." In the second two, the names of fictional characters is crucial; it is difficult to understand these quotes without knowing that Bilbo is a homebody of a hobbit, or that Japhy is a Zen poet seeking transcendence in the American West. (The fact that Bilbo's name is important can be seen in the fact that the user who posted this quote actually replaced "he" with it.) Likewise, the first quote, from Ende, is not just an excerpt from a narrative but is itself a meta-statement about the narrative, in which the worlds of the protagonist's life and a magical book have merged.

On a linguistic level, these unpopular quotations are more strongly "cohesive" (Halliday & Hasan, 2014) with their source text than typical popular quotations. Sometimes unpopular quotations are almost totally incomprehensible without some reference to the source text:

*There was me, that is Alex, and my three droogs, that is Pete, Georgie, and Dim, and we sat in the Korova Milkbar trying to make up our rassoodocks what to do with the evening. The Korova milkbar sold milk-plus, milk plus vellocet or synthemesc or drencrom, which is what we were drinking. This would sharpen you up and make you ready for a bit of the old ultra-violence.*

A reader unfamiliar with Burgess's *A Clockwork Orange* would have trouble parsing this passage. Then again, perhaps the same could be said for the first-time reader of this book; these are the first words of Burgess's novel, and they immediately barrage the reader with the idiolect of Alex and his dystopian gang ("milk-plus," "droogs," "rassoodocks"). The passage, and the quotation itself, are thus intentionally confusing.

It is not at all surprising that such quotes are not frequently repeated. Quotes that make general statements that are easily decontextualized from books by this very fact have an unbounded audience. A reader unfamiliar with Tolkien is not likely to be charmed by the repetition of a moment of Bilbo's homesickness. Yet a reader of *A Clockwork Orange* might very well understand and appreciate Alex's introduction to his life of futuristic gang violence. Such unpopular quotes are not necessarily deficient. Rather that they have a somewhat different

purpose—not to pull wisdom out of a story but to relive in miniature a particularly poignant narrative moment, one that was meaningful to the quoter. They are not just from novels; they are celebrations of these novels, designed not for unfamiliar readers but for fans. I call them "readerly" quotes.<sup>62</sup> They more directly rely on a reader having familiarity with the story, and in that sense they presume the "deep attention" of an audience—or, perhaps they do not expect an audience at all, counting only on their own attention.

### **Blog-Level Trends: Popular vs. Unpopular Blogs**

I hypothesized that book quote blogs that are consistently popular would less frequently use Person Names. Comparing BQ Pop Blog Corpus with BQ Unpop Blog Corpus, I found this to be the case. On average, blogs in the BQ Pop Blog Corpus contained 4% quotes with Person Names. Blogs in the BQ Unpop Blog Corpus contained 11% quotes with person names. This difference was statistically significant according to a two-tailed t-test ( $t=4.51$   $p<.001$ ). Blogs in the BQ Top Blog Corpus (average character length: 176.6) were also somewhat shorter than those in the BQ Unpop Blog Corpus (average character length: 224.2). This difference was significant according to a two-tailed t-test ( $t=5.09$ ,  $p<.001$ ) There were no differences in terms of frequency of imperatives or repetition.

I interpret this as evidence that popular and unpopular bloggers tend to blog different types of quotes, with popular blogs posting quotes that solicit the widest attention possible.

I do not mean to imply that popular blogs are more popular *because* they post fewer readerly quotes or shorter quotes. Further ethnographic analysis could shed light on the causes of these differences. While the present study cannot speculate regarding the motivation of popular versus unpopular quote bloggers, it can suggest that popular blogs do tend to eschew those quotations that would make quotations less "spreadable" (Jenkins, Ford, & Green, 2013). This invites more analysis of the ways that shaping one's media to be spreadable (i.e. to get the most attention) may be a skill or a literacy in its own right, one to which users have differing levels of access. It may also be the case that some users are "narrowcasters," intentionally catering to a smaller audience.

### **The Role of the Author**

Repetition, length, lexical choices, syntactic patterns—these surface features in various ways get at what makes quotations seem particularly evocative and worth repeating. However, quotations are (in general) not simply wise or beautiful sentiments. Most of them come with an authorial attribution. These words are by Jane Austen, those by John Grisham.

I now report the result of the sole "experimental" question given to quoters in the initial questionnaire. Quoters were shown a quotation by Toni Morrison or Fyodor Dostoevsky, either with or without authorial attribution, and were asked to rank its quotability on a Likert scale ( $n=115$  respondents to this question).

My hypothesis was that questionnaire respondents would rank those with the authorial attribution as more quotable, and this turned out to be the case. On average, unattributed quotes were ranked 5.5 (out of a possible 10), while attributed quotes were ranked 6.6. This difference was determined to be statistically significant according to a two-tailed t-test ( $t=2.48$ ,  $p<.05$ ).

This is a good reminder that readers do not evaluate the linguistic aspects of quotations in

---

<sup>62</sup> See Rowberry (2016) for a complementary discussion of similar phenomena in Kindle highlights.

a vacuum. Rather they see the beauty of words through their own conceptions of *who* is quotable.

### Are Quotations Network Specific?

So far this chapter has compared *quotations* to *non-quotations* as well as *popular* to *unpopular* quotations. While I have complemented quantitative analysis with a more fine-grained linguistic approach, this analysis is clearly rooted in other analyses of quotation as a genre (Bendersky & Smith, 2012; Danescu-Niculescu-Mizil, Gamon, & Dumais, 2011). However, are quotations simply a sort of transcendent genre, independent of the community that circulates them? Or is there something "Tumblr-specific" about Tumblr quotations?

As argued above, certain kinds of pronouns, including the singular first person pronoun "I," can often be used in ways that invite the reader of the quote to project herself into the quote, understanding it pseudo-autobiographically. These quotes are tailored to one of quoter's reasons for quoting, the expression (in a sometimes veiled, protected way) of their own intimate feelings. If there is something specifically Tumblr-esque about this purpose for quoting, then other corpora of quotations may have less of quotations containing "I," as well as other first person singular pronouns. Table 6 compares the ratio of such quotations in the Twitter Quotes Corpus, the BQ Corpus from Tumblr, and the GQ Corpus from Tumblr.

Table 6.

*Comparison of use of singular 1st person singular pronouns, Twitter vs. Tumblr.*

<b>Corpus</b>	<b>% 1st Person</b>
Twitter Quotes	9%
Book Quotes	29%
Quotes	37%

The BQ Corpus and GQ Corpus both contained at least 20% more first person pronouns than the Twitter Quotes Corpus. This is consistent with the hypothesis that speaking for the quoter as the intimate but nonspecific "I" is not just a general role of quotations but one that is especially important on Tumblr.

Is Tumblr indeed more "emotionally raw" than Twitter, as one quoter suggested (see Chapter 4)? Table 7 compares the percentages of quotes in the BQ Corpus containing either positive or negative words compared to the Twitter Quotes Corpus.

Table 7.

*Comparison of use of positive and negative words.*

<b>Corpus</b>	<b>% w/ Positive Word</b>	<b>% w/ Negative Word</b>
Twitter Quotes	53%	36%
Book Quotes (Tumblr)	48%	53%

By this measure, Twitter quotations are more positive than they are negative; this is not the case for Tumblr book quotes. The difference between these two distributions was statistically significant according to a chi-squared test ( $\chi^2=40.04$ ,  $p<.001$ ). While this measure is only a rough proxy for positive or negative sentiment, it does suggest that there are tonal differences between Twitter and Tumblr quotes. This finding is in turn consonant with my sense of the stylistic differences between these two types of quotes. Quotes shared on Twitter struck me as particularly anodyne—an admittedly subjective description:

*The miracle is this, the more we share, the more we have.* (Attrib. Leonard Nimoy)

*Don't let today's opportunities become tomorrow's what ifs.* (Attrib. Pat Flynn)

*What you choose today will determine who you are tomorrow.* (Attrib. Tim Fargo)

Taken together, these differences suggest that "quotations" are not a unified, homogenous genre. The fact that Tumblr quotations are characterized by frequent use of first-person pronouns and a willingness to discuss dark or negative emotions once again makes sense in light of the fact that quotations may be for some quoters a form of self-expression of intimate thoughts, including those that they may feel less inclined to express outside of the anonymity of the network.

### Discussion

This chapter has used the techniques of computational linguistics to analyze quotations, especially those from books. Building on a growing body of computational linguistic research on quotability (Bendersky & Smith, 2012; Danescu-Niculescu-Mizil, Cheng, Kleinberg, & Lee, 2012; Booten & Hearst, 2016; Rowberry, 2016), it has treated quotes as data to be cut to pieces, abstracted, tallied, compared, and calculated. It recounts not a single "experiment" but a recursive investigation of an admittedly nebulous concept—quotability. This "distant reading" has revealed aspects of quotations that would be difficult to notice by traditional hermeneutic means. The result is a coherent but by no means homogenous typology of quotations and, implicitly, quoters. Different quotations involve different strategies for repurposing the language of the other through appropriation (Bakhtin, 2010) and for making language "spreadable" (Jenkins, Ford, & Green, 2013, pp. 9).

On the most basic level, this analysis draws a distinction between those quotes—the majority of them—that require no knowledge of the quoted book (and are thus easily "decontextualized") and those that required some knowledge of the source. The former are both more common and more successful in that they are more likely to be repeated. Another distinction may be made between quotations that are long and those that are short. Once again, brevity is a generally a virtue in the world of Tumblr quotations, confirmation that they are designed for readers' "hyper" attention, not their "deep" attention.

Quotations are often not just language but what we might call "well-organized" language. They contain more repetition, signaling an overt poeticism. Yet they also tend to follow some recurring linguistic patterns, such as what I have called the "mini-volta." In this feature, a hairpin logical turn is itself a form of poetic power and beauty. Quotations, despite their brevity, can be almost spring-loaded with linguistic energy. Quoters themselves, as Chapter 3 demonstrated, often regard quotations as containing some powerful eloquence. Underlying the practice of quotation culture is a celebration of aesthetically powerful language.

Beyond these distinctions between "successful" and "unsuccessful" quotes are observations that come closer to describing what different roles quotations may play. Often they make sweeping, bold statements. They like to tell us what is never the case or what is true in all cases. They like to tackle the biggest concepts—life and love—without any sense that such concepts are too big to fit into one or two sentences. (The only thing diminutive about these texts is their size.) Such quotations are what Burke (1973) would call "equipment for living," providing people with strategies for naming, understanding, and even acting upon common situations. Yet quotes can also take on a somewhat more personal function, embodied by those that invite the reader to imagine themselves as the deictic "I," projecting the identity of someone else (perhaps a boyfriend or girlfriend) into the position of the deictic "you." The popularity of these quotations make sense in light of quoters' assertion that quotation can serve as instruments

for *self*-expression, providing the words that the quoter cannot generate spontaneously. Tumblr quotations are indeed well-formed for the particular rhetorical purposes of quoters. Furthermore, this use of quotations is in itself an interesting variation on what it means to "appropriate" the words of others. For Bakhtin (2010), appropriating another's language is often a great struggle, since some words "stubbornly resist" being put to new expressive purposes; these "vague pronoun" quotations are designed to invite the reader herself to easily identify with and thus appropriate. Just as a clothing designer creates a garment with shapes and openings that anticipate the body of many different wearers, the creators of "vague pronoun" quotations anticipate that different readers will be able to "slip inside" these quotes.

There are also those quotations—rare but present—that must in some way remind the reader of a particular book, perhaps bringing to mind a powerful scene or beloved character. Let us call these "readerly" quotations, since they seem to demand that one have read the source text in order to fully appreciate the quote. Though physically separate from the source text, they are not as truly "decontextualized" as other quotes. Compared to their popular counterparts, they are more aligned with "deep attention" both in that they themselves are often longer and more abstruse and in that they imply familiarity with the source text. Recall that in Chapter 3, quoters described the ways that some quotations may indeed be best appreciated with the context of the book in mind. While such quotations are not as popular as those that do not require understanding, both exist within Tumblr's ecosystem. There is not one type of quote.

In an interview, one reader-quoter compared the difference between popular and unpopular quote blogs as the difference between blockbuster movies and indie films. While quotes that are easily decontextualized dominate the textual economy of Tumblr, this does not mean that quotes that deviate from this norm are failures. The latter kind may be less popular, but their offerings may simply yield rarer pleasures for a smaller audience.

This study primarily examined book quotes from Tumblr, yet it also found some tantalizing differences between quotations on Twitter and those Tumblr quotes. At least in this comparison to Twitter, Tumblr quotes seem to be more negative, and they contained fewer first person words. These findings make sense in light of previous research on Tumblr's importance as a place for girls to express negative emotions (Shorey, 2015). Perhaps quoters on Twitter do not turn to quotations as tools for personal "self-expression." In any case, this study has suggested some of the ways that the kinds of texts that people share, and indeed the ways that they read, are inextricable from the larger purposes that undergird individuals' practices on social media.

## CHAPTER 6 IN CONCLUSION

Books are artifacts with a deep and abiding history that belong in and to our age—no more and no less so than flat-screen televisions, MP3 players, computers, and other so-called cutting-edge technologies (Striphas, 2009, pp. 4)

### **Summary of the Study**

This study has explored the ways that the destiny of print media, especially books, has become inseparable from one particular technology that seems, if no longer "cutting-edge," certainly contemporary: social media.

Chapter 3 explored the ways that the impossibly frenetic and diverse textual world of Tumblr is home to a rich culture of quotation. Often, though certainly not exclusively, quoters draw quotations from books as well as from discourses that not long ago were primarily accessed through print media, such as philosophy and literature. A main contribution of this chapter is to witness some of the sheer variety of such quotations, providing something of a partial typology of different types of quotes. These include the misquotation attributed to a well-known author, the quote that corrects this misquotation and makes clear the correct source (what I call "vernacular bibliography"), the photographic quote that luxuriates in the sumptuous pleasure of the printed word, and the "transmedia" quote that freely associates a literary text with its new media instantiations (especially film). Analysis of hashtags reveals other dimensions of quotation as a practice. Quoters often use tags that suggest that quoting can happen in the process of reading. In the case of philosophical quotes, they often use tags that seem drawn from academic discourse.

Chapter 4, putting aside most of this diversity, considered the phenomenon of the "#book quote" through the words of Tumblr users themselves. For them, most of whom are young women, quoting is a form of self-expression, and this self-expression is a form of self-care and validation-seeking. To share a quote on Tumblr is to place oneself into an imagined network of authors and readers who share the same emotions, even the same thoughts, though in a way that may allow for "plausible deniability" of what was expressed. According to quoters themselves, quotation may be a gendered practice in the sense that male Tumblr users may be less inclined to make public their emotions, even in an anonymized fashion beneath the protective veil of another's words. Authors such as John Green may be popular on the site exactly because their works are "quotable" in ways that align with young and mostly-female Tumblr users' expressive intentions. For quoters, quoting can at times connect to "deep attention" in interesting ways, most importantly by serving as "vernacular advertisements" for books (which they sometimes seek out and read). Quotation may even represent a "deep" way of paying attention to literature, since quotations are small enough to ponder and to store away in one's memory.

Chapter 5 used computational linguistic techniques to explore corpora of Tumblr quotations (mostly "#book quotes"). Analysis of "key" features revealed that quotations are in certain ways well-suited to the expressive purposes of quoters. For instance, they are often designed to be general yet highly emotional first person statements through which quoters may "voice" themselves. Analysis of other features reveals some of the ways that quotations are made to be "generalizable" as well as poetic. Comparison of popular and unpopular quotations from books reveals that popular quotations are designed to be more sharable. Unpopular quotations



tend to be longer and contain more person names. References to characters make a quotation less "generalizable," less easily understood outside of the textual context from which the quotation was plucked; often these quotations seem like they are designed to be enjoyed by readers who are already very familiar with the source text. Quoters themselves expect different sorts of "attention."

### **Limitations of the Study**

- Phase 1, focusing on the textual ecosystem of Tumblr, observed many of the complicated ways in which the quotation culture of Tumblr intersects with old media. The remainder of the dissertation took a narrower view by mostly focusing on "book quotes" of the quote-type and those who share them. Those who share philosophical quotes or who create "fan edits" that seem to be quotes may have different perspectives on the goals of quotation.
- This study used certain computational techniques to make certain claims about the makeup of quoters on Tumblr. (For instance, Chapter 3 observed that certain quoters seem more "bookish" than others.) More sophisticated computational analysis (e.g. Ikeda et al., 2013) could have perhaps revealed a more complex picture of the communities that exist on Tumblr. For instance, while John Green is clearly a very popular author whose works are quoted on Tumblr, perhaps there is a community that is devoted to his work (and other YA novels) and another community that eschews these genre in favor of more "literary" fare (e.g. Murakami).
- Those who answered my questionnaire are not guaranteed to be representative of book quoters as a population. This is due to the fact that I found the need to sample them non-randomly (as described in Chapter 2) and to the possibility of selection bias (those who answered my questionnaire may be different than those who did not).
- Studying particular practices on social media, the question remains to what degree certain phenomena are simply local instantiations of "general" trends. For instance, I observed that those who share book quotes may be somewhat more "bookish" than those who reblog them. Indeed, this is not to say that this is not generally true on Tumblr; people who post fashion photos may likewise be more "fashionable" than those who reblog them. A different design might have done more to compare the dynamics of bookish culture within Tumblr "in general."
- This was broadly a case study of Tumblr, though in certain ways the study thought comparatively about Tumblr's relationship to other networks such as Goodreads and Twitter, mostly through the words of quoters themselves, as well as through Chapter 5's brief comparison between Twitter and Tumblr quotes. However, in general the study was not designed to be fully comparative in a way that would shed light on the ways that quotation culture on Tumblr is or is not different than the kind that occurs on other networks. Further quantitative as well as ethnographic research would be needed to rigorously explore the ways that quotation on Tumblr is/is not different than on other sites (and indeed in other non-digital contexts).

### **Ethnography and Computation**

This study is an example of a genre that I am calling "Computational Textual Ethnography." While ethnographic in intent and perspective, it made use of computational techniques at all levels, from targeting Tumblr users as possible research participants to analyzing the linguistic nature of quotes.

It is now common to speak of "big data," as if the only thing at stake in the use of computational research is the number of data-points and scale of the phenomena observed. Certainly this study has at moments turned its gaze to relatively "big" sets of data, at least compared to other ethnographies of online practices, communities, and texts. And, yes, this has allowed me to observe some large-scale trends: for instance, in the kinds of quotations that quoters share. But, as ethnography, this study has also at times focused on small pieces of data, from a particularly fascinating text to a perhaps-idiosyncratic but illuminative perspective shared by a research participant.

It is easier for research to stay at one scale. Viewing the world from above, as "trends" that can be observed statistically, one need not reckon with (or even observe) the teeming underlife of contradictions that lurks on the forest floor of one's data. (A piece of data that "goes against the trend" is no less real than one that does not.) Likewise, staying close to a few examples, a few "telling cases" (Mitchell, 1984), relieves one of the responsibility of situating these cases in a larger constellation of phenomena. The integration of quantitative methods in research with an ethnographic perspective offers a potential way to work through this double bind. Still, this methodology could no doubt be refined and extended in various ways, especially by making it more participatory. During this research my use of computational techniques was informed and inflected by my encounters with quoters on Tumblr, but I still carried out these analyses on my own. For instance, during Phase 1 of research I chose how to visualize the relationship between hashtags, while in Phase 3 I chose the linguistic features of quotations to analyze. Future research in the vein of "computational textual ethnography" might invite participants within a textual ecosystem to help guide the design and production of quantitative texts—e.g. visualizations or even simple tabulations—in order that these text might better reflect their own concerns and curiosities.

### **Implications**

#### **Beyond Hyper and Deep**

This study began by voicing a distinction between "deep" attention—the kind of attention solicited by engagement with traditional, long-form media, especially books—and "hyper" attention—the kind of attention solicited by digital and perhaps especially social media.

Perhaps the most significant take-away from this study is that this binary typology is simply insufficient. The space where the book meets social media is an unstable borderlands where users/readers are devising a multiplicity of practices that link old and new media. Within this culture, the quotation is an ambiguous artifact. Certainly it can be a mere snack, a morsel to be quickly consumed from one's endless feed, not fundamentally different from a photo of an adorable cat short looping GIF of a surfer riding a wave. As Evens (2012) argues, Web 2.0 media lends itself to this flatness in which the distinctions between all forms of media are rendered functionally equivalent, subject to the same operations of sharing.

While this is a true part, this is only part of the story. "Deep" attention coexists in some cases happily with "hyper" attention; stumbling upon a perfect quote can end there, or it can spur the reader to seek out the book and read it. Those who worry that reading is a dying art should find consolation in this fact. The book, derided as hopelessly outmoded, turns out to have found a way to use digital media to help ensure its continued survival and relevance.

Yet this is not just a story of "hyper" vs. "deep." This is also a story of different ways of paying attention. A quotation may be a general, decontextualizable sentiment, or it may require readerly knowledge of the source text. A quotation may be shared with little regard for its provenance or veracity, or it may be meticulously annotated with bibliographic information or

even linked to a reliable source. A quoter may quote either academically, constructing a collection of quotes that might well be drawn straight from a syllabus, or promiscuously, juxtaposing philosophy and literature with a wide variety of other media and genres. A quote may be discovered on a blog forgotten instantly; it may also become ever-after a familiar object in the reader's memory. All of these tactics, these micro-practices, represent different ways of paying attention to literature

In some sense, this diversity makes for an unsatisfying conclusion. Extreme beliefs (such as that young people messing around on the internet spell the death of literacy, or that reading is being transformed for the better) make for better headlines. But perhaps what is really needed in our conversations about "the fate of the book" is simply a broader and more textured collection of accounts of individual readers who make use of new and old media together. What is clear from this study is that readers need not be puppets who rehearse the logics demanded by the media that they consume. Rather it is individuals themselves who, in the fabric over their own lives, weave together different ways of "paying attention" in order to meet their goals.

### "Social" Reading

Still, amidst this diversity, there are clear trends. Perhaps the most important of these is the way that quotations tend to be used in the process of self-expression and self/other-care in ways that are highly "social."

Certainly quoters' forms of quoting are distinct from at least an idealized version of traditional reading. For Birkerts (2006), reading a novel in particular is an experience characterized by its intense privacy. It is by entering into the world of the novelist's creation, that the reader's "soul"—that "consistent condensation of the self"—becomes "more present": "[t]he energies that otherwise tend to stream outward through a thousand channels of distraction are marshaled by the cadences of the prose" (pp. 85). In our hyperconnected world, characterized more and more by "unitary life," a book is one of the few places where one can really be alone (although, ironically, that aloneness is facilitated by the written words of another person).

For Birkerts, digital technology is fundamentally opposed to this precious aloneness, and he prophesies that at some point in the future "we will conduct our public and private lives within networks so dense" that "it will make almost no sense to speak of the differentiations of subjective individualism" (pp. 131). Likewise, Evens (2012) argues that social media technologies necessarily lead to an ideology that attempts to erode the subjective boundaries between discrete individuals:

On Web 2.0, alienation has become impossible, since each subject discovers herself there, her *simpatico* guaranteed by the fact of the encounter, which inscribes itself on the Web.

Does the increasing insinuation of digital and social media make impossible the salutary aloneness of reading? Chapter 2 revealed that the major reason for quoting on Tumblr is to feel less alone, to recognize one's own experiences and emotions as universal or at least widely shared. Of course, in some sense reading has always been social in different ways at different times; discussing a book in a classroom or for that matter in a book club likewise might be said to impinge upon the privacy of the reading experience. But what goes on on Tumblr is often not the discussion of books as literary objects but the recognition of sentiments as "resonant" with one's own emotions. The goal of quoting is in some sense to make sure that "alienation has become impossible."

At the same time, this practice occurs *simultaneously* with traditional reading. Birkerts

might be able to accuse the quoter who quotes while reading of violating the sanctity of readerly isolation, of turning outward just at that moment when one should turn inward. Yet what about the quoter who does not do this exactly but instead reads a novel first before deciding to post some quotes to Tumblr? In this case, perhaps even Birkerts would agree that nothing has been "lost" exactly (even if he were to doubt that something really has been gained through later making reading "social"). Just as quoting straddles the boundary between different forms of attention, it also does not *demand* any particular form of subjectivity—whether that of the isolated individual or the digitally-connected league of simpatico strangers (though it clearly favors the latter). In general, the study of real readers suggests that a book can have different uses to even a single reader at different moments.

One more finding from the previous chapters bears repeating here: even as Tumblr provides a place to become un-alienated, it is also a place where old, even solitary modes of reading are celebrated. This is seen most readily in those quotes which are photographs of books (and indeed in the larger trend of book photography). Such images do the paradoxical work of making public the private experience of reading while also fetishizing it; one may visit Tumblr only to see isolated, print-based reading repackaged as a beautiful, sumptuous leisure. The network advertises experiences that it cannot itself provide. Perhaps especially for Tumblr users, who are indeed so connected, the aloneness of print takes on an even greater allure simply because it now seems so much rarer.

### **Beyond Bibliophilia**

The nascent body of research literature on the intersection of social media and reading tends to emphasize the ways that book-centered websites like Goodreads and LibraryThing provide spaces for readers to connect with other readers and to share recommendations for books (Naik, 2012; Nakamura, 2013; Vlieghe, Muls, & Rutten, 2016). This is true of Tumblr as well. Even though Tumblr is not a book-focused social network, some users are certainly on Tumblr for many of these same purposes. In fact, this study has suggested that for at least some bibliophilic social media users, Tumblr plays a similar role. This network may have its specific affordances (especially in its facilitation of serendipitous discovery), and readers may combine it with other networks, especially Goodreads. Stepping back for a moment to consider Tumblr as one network among many, the fact that users use Tumblr as a sort of unsanctioned Goodreads attests to the imagination of social media users and the passion of readers. Companies such as Goodreads are no doubt valuable in the ways that they encourage and facilitate intersections between social media and reading. Still, they are hardly necessary for the cultivation of a vibrant culture of reading online.

However, Tumblr is also interesting precisely because it is a general-purpose social network, not an explicitly bookish one. Tumblr is home to readers—and it is also home to fans of a variety of other media. Tumblr is filled with quotes from books—but also GIFs of Rihanna, humorous memes, and photographs of perfectly-posed food. On Tumblr, one can be overtly bookish, posting only book quotes (and pictures of books, and reviews, and other bibliophilic content). But one can just as easily reblog a quote from *Jane Eyre* one minute and picture of a vintage purse the next; books may be just another type of text in an infinite tapestry of curious and captivating media. As Chapter 3 demonstrated, those who consume book quotes may be less self-consciously bookish than those who post them. One might argue that a site like Tumblr is actually a superior site for the cultivation of bookish culture because these books, in the form of quotes, likely reach social media users who do not think of themselves as bibliophiles. On

Tumblr, books break out from the book club.

The quotation itself is an artifact that straddles the boundary between bibliophilic culture and internet culture more broadly. It serves bibliophilic purposes like the discovery of new books. And they may at times expect a "readerly" attention, serving as reminders of characters or fictional moments. But more often they are also *generalizable*, making them easier to appreciate without any special knowledge of the source text. And facilitating self-expression and self-care are more general and more fundamental purposes than facilitating bookish discourse.

### Lingering and Emergent Questions

This study has carved a window into Tumblr, answering some questions while also generating others. Below I list a few questions that emerged in the course of the study that could be answered by further research as well as some possible next-steps in the study of contemporary literacy on the web:

- **Trajectories of Learning.** This study generally did not consider the ways that individuals as well as communities and these communities' practices may change over time. As recent approaches in the study have revealed, there is much to be gained from the study of individuals' "learning lives" (Erstad, Gilje, Sefton-Green, & Vasbø, 2009). For educators especially it would be useful to know more about the paths of peoples' relationships to books *before* they started blogs with names like "mymostbelovedbooks." For instance, did the people who engage in online bibliophilic culture usually have positive experiences of literature in school contexts, or is quoting books on Tumblr an outright rebellion against school-based regimes of literacy? Looking at "learning lives" could also suggest ways that readers evolve over time. In Chapter 4 I briefly recounted the words of a quoter who went through what might be called her "John Green phase." After discovering this author through quoting, she later grew weary of him. Studying who quotes quote and what sorts of quotes they produce over time could shed some light into the ways that literary taste changes. (It would be fascinating to know at what age one is most likely to quote Shakespeare, Charles Bukowski, or John Green.)
- **Authoring for Virality.** As briefly mentioned in the study, John Green's popularity on Tumblr may be due in part to his active presence on the site but also to a certain inherent "quotability" in his work. Other writers are arguably forging even stronger connections between literary authorship and social media presence. A 2015 *New York Times* article (Alter, 2015) considered the rise of what it called a "Web Poet's Society," a small cadre of writers whose work first found an audience on social networks, including Tumblr and Instagram. In fact, their works are *designed* for these networks. The poems of Tyler Knott Gregerson, for instance, are often "brief meditations of love and longing." In other words, they are a lot like quotations in general, except they *begin* as short texts. They are "pre-quoted." Indeed, on Tumblr many users produce and share literary texts that are defined by their brevity (Booten, Freedman, & Hull, 2017). What sets authors like Gregerson apart, however, is that their web-native literature, perfectly attuned for hyper-attention, has made its way into print. (Several of the quoters in Chapter 4 described discovering similar books through Tumblr quotations, including published works by Lang Leav and Rupi Kaur). Literacy (and literary) studies should continue to explore the ways that literary production itself is changing to become more "sharable" in today's economies of attention. In this respect, a sociological approach to contemporary literature could proceed hand in hand with the study of other media that may have grown in popularity via social media, including film

and television, and the study of social media celebrity more generally (see Marwick & boyd, 2011).

- **Literary Taste and Cultural Capital.** As Bourdieu's work on "cultural capital" and demonstrated, literary taste is a projection of class distinctions (Bourdieu, 1984, 2011). Let us assume, as seems reasonable, that carrying around a copy of Hegel's *Aesthetics* or Berryman's *Dream Songs* once served as a fairly strong visual clue that one has had access to a certain elite form of education. The ability to talk about them in a class or at a party represents the successful manifestation of a class position as a way of being. Compare this to the example given at the end of Chapter 3: a quotation from Heidegger, originating in a blog filled almost exclusively of literary and philosophical texts, promiscuously reproduces itself (via "reblog") in various contexts that are not themselves overtly literary in their disposition. It would not be surprising if more people in the past year have read Heidegger on various social networks than have held one of his print books. One can view this optimistically; philosophy and "high" literary texts are now the property of everyday social media users of all stripes! Likewise, if John Green is approximately as popular as Shakespeare on Tumblr, then perhaps the distinction between "high" and "low" forms of literature are themselves becoming less stable. ("Omnivorous" taste, rather than outright snobbery, has been associated with "high-status persons" [Peterson, 1992]) However, it is impossible to make these claims without really considering class distinctions directly. To put it bluntly: what are the class affiliations and educational experiences of book quoters online, and how do these correlate with different ways of experiencing literature in the context of digital media?
- **Quoting and Commerce.** In her discussion of Goodreads, Nakamura (2013) is suspicious of the ways that the readerly labor of book lovers on that site ends up "in the service of commerce" (pp. 244), and indeed the media that Tumblr users share—sometimes for deeply personal reasons—have the added effect of making Tumblr itself money in the form of advertising. As this study has also suggested, social networks are not free from the influence traditional organs of publishing, including agencies and presses. The rise of new forms of vernacular advertisement, including quoting, no doubt presents a threat to the traditional strategies the publishing industry has used to acquire the attention and dollars of readers. Yet it perhaps also represents for them an opportunity to expertly target readers as well as bibliophilic communities with particular books. Future studies could directly consider the ways that publishing companies and other media industries intervene in social media platforms.

### Interdisciplinary Connections

This study may be of some use to scholars and practitioners in other fields. I close with two different reflections on the interdisciplinary implications that may be drawn from this study. The first ponders Tumblr quoting's position in the history of reading and the book. The second suggests how knowledge of quotation practices on Tumblr may be useful for educators in formal instructional settings.

#### Quoting on Tumblr: Is "New media" new?

The title of this dissertation itself implies a disjuncture between some overarching regime of "old" media and new "social" media. Yet this study, not historical in dissertation, has left open the possibility that while Tumblr itself is new the digital practices it facilitates, including and

perhaps especially the collection of quotations, is not. I return to the question of the ways that quoting is or is not like "commonplacing."

Certain comparisons are obvious and not particularly helpful. That Tumblr is not awash in either the wisdom of church fathers or classical poetry is not at all surprising in light of the fact that its core users are neither medieval monks nor Renaissance schoolboys. Yet other differences help illuminate important cultural aspects of book quoting on Tumblr.

**Scholarly vs. Vernacular Practice.** First, Tumblr quoting is unlike certain versions of the commonplace book in that it is disconnected from the arts of rhetoric or composition more broadly. Starting in antiquity, the commonplace book was a goal-oriented object, not an end in itself. According to Cicero, quotations held within them could provide ammunition for the rhetorician; a few choice phrases could be mobilized on-the-fly and woven into extemporaneous argument. During the Early Modern period, commonplace books provided information and stylistic examples for writers, and the writing of authors such as Montaigne (rich as it is with aphorisms) demonstrates the way that reading could be the first step in writing (Allan, 2010).<sup>63</sup> Behind this use of the commonplace were theories of the nature of composition. Earlier I abbreviated Seneca's wisdom on a bee-like method of reading; the second half suggests that the keepers of commonplace must "blend" the collected textual fragments in their own compositions, creating a text that is new yet "betrays its origin." (Proponents of "remix" aesthetics [and the related pedagogy] have Seneca on their side.) While Tumblr quoters sometimes described oblique ways of using quotes as "inspiration" of one form or another, these examples were relatively rare; for most of the quoters I interviewed, gathering fodder for writing (and certainly not *academic* writing) is not a major purpose of commonplacing.

Today's quoting on Tumblr takes place as a vernacular practice by readers, mostly in the course of their pleasure reading. During the Renaissance, however, keeping a commonplace book was a foundational part of humanist pedagogy. Distinguishing commonplace books from other miscellaneous book-forms, Moss (1996) lays out the pedagogical importance of the commonplace book in Renaissance education:

[E]very Latin-literate individual started to compose a commonplace book as soon as he could read and write reasonably. It was formative and it was programmatic. It shaped the way he thought and determined the way he handled language.

It is hard to imagine an activity that is quite so universal in today's educational systems—except, perhaps, the standardized test. Commonplacing, as an official practice, was also the subject of intense theorization, something that predated and extended beyond the Renaissance. Both Erasmus and John Locke, wrote pedagogical treatises discussing how best to maintain a commonplace book. As Allan (2010) notes of the latter, it dramatically influenced how Georgian-era readers took notes while also spawning a series of competitors. Insofar as keeping quotations on social networks is indeed a form of commonplacing, then certainly commonplacing is as popular as it has ever been (in terms of total number of participants). Yet where is Tumblr's Erasmus? Quoting on Tumblr occurs on its own, disconnected from both formal schooling and mainstream discourses that describe the "right" way to do something as intimate as take notes.

There are exceptions to be made to these claims of fundamental difference between

---

<sup>63</sup>Commonplaces were also useful in professional fields, such as the law (Havens & Mooney, 2001), and in the burgeoning "Natural Science" (Blair, 1992).

quoting on Tumblr and commonplacing, some of them the result of the flexibility of commonplacing itself. As Moss (1996) notes, in the late 17th century the term "commonplace book" began to blur into less structured and less pedagogically-connected forms of private, miscellaneous note-taking. Making an argument for a somewhat more narrow definition of commonplacing, Allan (2010) argues that scholars themselves have been too quick to call any miscellaneous notebook a commonplace book. Without attempting to resolve any debates about what a commonplace book is, suffice it to say that Tumblr blogs—containing as they often do quotes, photographs of food, personal musings, and animated gifs—tend to be more miscellaneous than commonplace books *typically* were.

**Self-Care.** For quoters on Tumblr, quoting is connected in various ways to self-care. Quotes themselves may provide inspiration or guidance through a time of struggle, but they may also express the quoter's feelings in a way that makes her feel less alone and in fact "validated." The same very public form of validation could not be satisfied in the same way within the comparatively lonely confines of a private book. On the other hand, reading has always been social in certain ways, and sometimes in ways that facilitate self-care. In this aspect, quoting on Tumblr looks less like commonplacing than even earlier textual practices.

According to Foucault (2005), letter writing but also the collection of notes from reading played an important role in ancient stoic philosophy. The Greeks called these notes "hupommēnata," a word that signals that they were designed to facilitate the memorization of text. Once memorized, certain truthful philosophical statements or discourses ("logoi") would become one's "paraskeue"—that is to say, "equipment" that would be immediately available ("prokheiron", "ready to hand") to provide assistance and guidance when one is confronted with one of life's many trials. Perhaps there is an echo here of the ways that quoters, having memorized quotations, recall them frequently when confronted with interpersonal or financial difficulties.

Yet even in the ancient Greek context quotations could be shared. Sharing such "hupomnēmata," or integrating their contents into letters, was a way of taking care of another person as well as one's self, since in the context of this textual practice "[t]he advice you give to another is equally given to yourself" (pp. 361). This too is not so far from the ways that sharing quotations on Tumblr may at times be a form of self-care but also other-care.

There are also clear differences between this ancient practice and Tumblr quoting. *Hupomnēmata* would contain philosophical "logoi," not angsty *cris de coeur*. In some ways, the sharing of such statements is more like what Foucault identifies as later confessional practices, in which the subject must speak the truth about himself. But, if Tumblr is confession, it is a very strange sort in which one can reveal the depths of one's soul without vouchsafing a single personal detail, without uttering a word of one's own at all. And, insofar as digital quotation's status as self-expression is a particularly female practice, this further distinguishes it from its ancient counterparts.

### **Educational Implications**

This study has not been about *education* but rather *literacy*. However, the two are intertwined. Insofar as many Tumblr users are school-aged, quoting on Tumblr might be considered an "out-of-school literacy." Within the broader field of education/literacy studies, documenting and exploring young people's out-of-school literacies is itself a tactical maneuver that pushes back against ideas of young people as literate only insofar as they acquire the norms



of academic literacies (Knobel & Lankshear, 2003). There also exists a rich body of research that explores the potentialities and perils of trying to "bridge" young people's in-school and out-of-school literacies (see Hull & Schultz, 2002; Ito et al., 2013). This research also has a practical goal: any study of the ways that young people engage in literacy practices outside of schools and on social networks produces knowledge that may one day be "operationalized" by educators who aim to design more effective and more humane forms of instruction by capitalizing upon learners' rich experiences consuming and producing texts of all sorts beyond the classroom walls. I conclude by taking stock of aspects of digital quotation that may be especially important to educators.

**Possible connections between quoting and school-based literacy.** On the most basic level, this study has presented a somewhat optimistic account of the fate of print media in the context of social media. Insofar as digital media has become an important space for the cultivation and appreciation of books (even "academic" ones), teachers may be able to invent ways to connect their classrooms not just to Tumblr but to other online spaces of bibliophilia as well.<sup>64</sup> More specifically, this study has drawn attention to some of the ways that "deep" attention to traditional print media can actually be supported (rather than obviated) by the use of social media. Teachers may come to think of social media use, or certain forms of it, as a way of fostering exactly the sort of print-based literacies that still enjoy pride of place in most levels of formal educational instruction.

*Critical versus "uncritical" reading.* However, while the divide between "hyper" and "deep" attention may be less severe than sometimes imagined, this does not mean that there is not the potential for fundamental incompatibilities between the ways that books function in school settings versus online, interest-based social networks. As I claimed at the end of Chapters 3 and 4, quoters generally evince an "uncritical" perspective toward their quotations; rather they quote sentiments that "resonate" with them, that inspire them, that say what they cannot. From an academic perspective, to say that quoters are "uncritical" would be an insult, since criticality regarded as perhaps the most cardinal scholarly virtue. And educational research has long argued for the importance of helping students to develop capacity for "critical thinking" (see Ennis, 1962; King & Kitchener, 1994) as well as politically-attuned "critical literacy" (Shor, 1999) and "critical media literacy" (Alverman & Halgood, 2000).

I do not mean to denigrate the value of "criticality" of any sort. Indeed, in this age of "fake news," the proliferation of misquotations can be seen as an epiphenomenon of a larger lack of critical awareness regarding textual provenance and veracity on social media. As Felski (2011) has observed, however, "critical reading" has become an ever-present buzzword in higher education, yet there is a growing sense among humanists that the valorizing of this mode of reading has itself become uncritical. Critical reading implies not just dispassion toward but also suspicion of literary texts; to read something "critically" is largely to discover what is wrong with it, perhaps to uncover hidden ideological "symptoms" whose troubling existence were a foregone conclusion. As an alternative, Felski suggests that literary critics must reconnect themselves to the ways that non-critical readers approach understand but also make use of texts. The first use of literature described by Felski is one of the most important uses of quotation; to provide occasion for the reader to see herself, to "recognize" herself, in a fictional circumstance.

---

<sup>64</sup> See Vlieghe and Rutten (2016) for an example of using Goodreads with pre-service teachers.

Yet it is not just that quoters read uncritically; they may do so for the purpose of self-expression that is at once public and intimate. For some quoters, Tumblr is a kind of "safe space" for expressing, through the borrowed utterance of quotation, a deeply felt emotion that they might not feel confident sharing outside of their (perhaps anonymized) blog. As educators work to think creatively about the ways that school and out-of-school, online literacy practices can be usefully connected to school settings, they should be wary of the ways that online practices depend upon a sense of anonymity and freedom may be antithetical to the very notion of formal schooling.

*Towards a New Rhetoric.* On a more positive note, this study of quotation practices has also drawn attention to the ways that the humble written (or rather typed) word can still be captivating, even in a social network that is filled with all manner of vibrant images that instantly lure and dazzle the eye. Here, perhaps, the connection between quoting and commonplacing could reassert itself as part of contemporary education. As explained above, the commonplace book once served a pedagogical purpose by furnishing examples of powerful language that might be used for persuasive argument as well as powerful writing. What would it look like for educators to consider the new challenges of writing in the context of an economy of attention in which the eyes of readers cannot be expected but instead must be quickly seduced? A new rhetorical education, one tailored to the new condition of writing in the context of social media, would need to teach students about what it means to tailor language to capture the attention of online audiences. It might borrow from the ancient arts of rhetoric while also updating its linguistic and formal vocabulary to better describe new genres. And, above all, it might consider the ways that attention can be solicited carefully and responsibly, with regard for the ways that one's readers may find a text to be useful. In this respect, quoters—with their often quite-earnest belief that even the smallest text can lead to a dramatic change in someone's life—may serve as exemplars of a way of engaging in today's attention economies that is caring rather than crass, ethical rather than exploitative.

## REFERENCES

- Aarts, B. (1996). The rhetorical adverb simply in present-day British English. In C. Percy & C.F. Meyer & I. Lancashire (Eds.), *Synchronic corpus linguistics* (pp. 59-68). Amsterdam, The Netherlands: Rodopi.
- Allan, D. (2010). *Commonplace books and reading in Georgian England*. Cambridge, UK: Cambridge University Press.
- Alter, A. (2015, November 8). Web poets' society: New breed succeeds in taking verse viral. *New York Times*, p. A1.
- Alvermann, D. E., & Hagood, M. C. (2000). Critical media literacy: Research, theory, and practice in "New Times". *The Journal of Educational Research*, 93(3), 193-205.
- Androutsopoulos, J. (2008). Potentials and limitations of discourse-centred online ethnography. *Language@ Internet*, 5(8).
- Androutsopoulos, J. (2011). From variation to heteroglossia in the study of computer-mediated discourse. In C. Thurlow & K. Mroczek (Eds.), *Digital Discourse: Language in the New Media* (pp. 277-298). New York, NY: Oxford University Press.
- Baker, P. (2006). *Using corpora in discourse analysis*. London, UK: A&C Black.
- Baker, P. & McEnry, T. (Eds.). (2015). *Corpora and discourse studies: integrating discourse and corpora*. New York, NY: Palgrave Macmillan.
- Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics*. (Caryl Emerson, Trans.). Minneapolis: U of Minnesota Press.
- Bakhtin, M. M. (2010). *The dialogic imagination: Four essays*. M. Holquist (Ed.), (C. Emerson & M. Holquist, Trans.). Austin: TX: University of Texas Press.
- Barton, D. (2015). Tagging on flickr as a social practice. In R.H. Jones, A. Chik, & C.A. Hafner (Eds.), *Discourse and digital practices: Doing discourse analysis in the digital age* (pp. 48-65). Abingdon, UK: Routledge.
- Bendersky, M., & Smith, D. A. (2012). A dictionary of wisdom and wit: Learning to extract quotable phrases. In *Proceedings of the nAACL-hLT 2012 workshop on computational linguistics for literature* (pp. 69-77).
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge, UK: Cambridge University Press.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media, Inc.
- Birkerts, S. (2006). *The Gutenberg elegies: The fate of reading in an electronic age*. New York, NY: Faber and Faber.
- Black, R. W. (2005). Access and affiliation: The literacy and composition practices of English-language learners in an online fanfiction community. *Journal of Adolescent & Adult Literacy*, 49(2), 118-128.
- Black, R. W. (2007). Digital design. *A New Literacies Sampler*, 29, 115.
- Black, R. W. (2009). Online fan fiction, global identities, and imagination. *Research in the Teaching of English*, 397-425.
- Blair, A. (1992). Humanist methods in natural philosophy: The commonplace book. *Journal of the History of Ideas*, 541-551.
- Bolter, J. D. (1996). Ekphrasis, virtual reality, and the future of writing. In G. Nunberg (Ed.), *The future of the book* (pp. 253-272). Berkeley, CA: University of California Press.
- Booten, K., & Hearst, M. A. (2016). Patterns of wisdom: Discourse-level style in multi-sentence

- quotations. In *Proceedings of NAACL-HLT* (pp. 1139–1144).
- Booten, K., Freedman, S. W., & Hull, G. (2017). The internet's concept of story. *English language arts research and teaching: revisiting and extending Arthur Applebee's contributions*. New York, NY: Routledge.
- Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.
- Bourdieu, P. (2005). Habitus. In J. Hillier & E. Rooksby (Eds.), *Habitus: a sense of place* (pp. 43–52). Aldershot, UK: Ashgate Publishing.
- Bourdieu, P. (2011). The forms of capital. (1986). In I. Szeman & T. Kaposy (Eds.), *Cultural theory: an anthology* (pp. 81–93). Hoboken, NJ: Wiley-Blackwell.
- Boyarin, J. (1989). Voices around the text: The ethnography of reading at Mesivta Tifereth Jerusalem. *Cultural Anthropology*, 4(4), 399–421.
- Brandt, D. (2001). *Literacy in American lives*. Cambridge, UK: Cambridge University Press.
- Brandt, D., & Clinton, K. (2002). Limits of the local: Expanding perspectives on literacy as a social practice. *Journal of Literacy Research*, 34(3), 337–356.
- Brasel, S. A., & Gips, J. (2011). Media multitasking behavior: Concurrent television and computer usage. *Cyberpsychology, Behavior, and Social Networking*, 14(9), 527–534.
- Buckingham, D. (2007). Digital media literacies: Rethinking media education in the age of the internet. *Research in Comparative and International Education*, 2(1), 43–55.
- Burke, K. (1973). Literature as equipment for living. In *The philosophy of literary form* (pp. 293–304). Berkeley, CA: University of California Press.
- Burton, G. (2015). Silva rhetoricae: The forest of rhetoric. Retrieved from <http://www.rhetoric.byu.edu/forest.htm>.
- Camus, A. (1967). *Lyrical and critical essays* (E.C. Kennedy, trans.). New York: Knopf.
- Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. New York: WW Norton & Company.
- Chang, Y., Tang, L., Inagaki, Y., & Liu, Y. (2014). What is Tumblr: A statistical overview and comparison. *ACM SIGKDD Explorations Newsletter*, 16(1), 21–29.
- Chatman, S. B. (1980). *Story and discourse: Narrative structure in fiction and film*. Ithaca, NY: Cornell University Press.
- Collins, J. (2010). *Bring on the books for everybody: How literary culture became popular culture*. Durham, NC: Duke University Press.
- Corbett, E. P. (1965). *Classical rhetoric*. JSTOR.
- Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonçalves, M. A., & Benevenuto, F. (2011, June). Analyzing the dynamic evolution of hashtags on twitter: a language-based approach. In *Proceedings of the Workshop on Languages in Social Media* (pp. 58–65). Association for Computational Linguistics.
- Danescu-Niculescu-Mizil, C., Cheng, J., Kleinberg, J., & Lee, L. (2012). You had me at hello: How phrasing affects memorability. In *Proceedings of the 50th annual meeting of the association for computational linguistics: Long papers-volume 1* (pp. 892–901). Association for Computational Linguistics.
- Danescu-Niculescu-Mizil, C., Gamon, M., & Dumais, S. (2011). Mark my words!: Linguistic style accommodation in social media. In *Proceedings of the 20th international conference on world wide web* (pp. 745–754). ACM.
- Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). No country for old members: User lifecycle and linguistic change in online communities. In

- Proceedings of the 22nd international conference on world wide web* (pp. 307–318). ACM.
- Dawkins, R. (2016). *The selfish gene*. New York, NY: Oxford University Press.
- De Beaugrande, R. (1980). *Text, discourse, and process: Toward a multidisciplinary science of texts* (Vol. 4). New York, NY: Ablex Publishing.
- De Certeau, M. (1984). *The practice of everyday life*. (S. Rendall, Trans.) (Vol. 1). Minneapolis, MN: U of Minnesota Press.
- De Kosnik, A. (2016). *Rogue archives: Digital cultural memory and media fandom*. Cambridge, MA: MIT Press.
- Dijk, T. A. van (Ed.). (1985). *Discourse and communication: New approaches to the analysis of mass media discourse and communication* (Vol. 10). Berlin, Germany: Walter de Gruyter.
- Dowling, D. (2014). Escaping the shallows: Deep reading's revival in the digital age. *DHQ: Digital Humanities Quarterly*, 8(2).
- Dubremetz, M., & Nivre, J. (2015). Rhetorical figure detection: The case of chiasmus. *Fifth Workshop on Computational Linguistics for Literature*, 23.
- Ducheneaut, N., Yee, N., & Bellotti, V. (2010). The best of both (virtual) worlds: Using ethnography and computational tools to study online behavior. In *Ethnographic praxis in industry conference proceedings* (Vol. 2010) (pp. 136–148). Wiley Online Library.
- Elliott, R., & Timulak, L. (2005). Descriptive and interpretive approaches to qualitative research. In *A handbook of research methods for clinical and health psychology* (pp. 147-159). New York, NY: Oxford University Press.
- Erstad, O., Gilje, Ø., Sefton-Green, J., & Vasbø, K. (2009). Exploring "learning lives": Community, identity, literacy and meaning. *Literacy*, 43(2), 100–106.
- Ennis, R. H. (1962). A concept of critical thinking. *Harvard Educational Review*, 59(3), 297-394.
- Evens, A. (2012). Web 2.0 and the ontology of the digital. *DHQ: Digital Humanities Quarterly*, 6(2). Retrieved from <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>.
- Fairclough, N. (1993). Critical discourse analysis and the marketization of public discourse: The universities. *Discourse & Society*, 4(2), 133–168.
- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. London, UK: Routledge.
- Felski, R. (2011). *Uses of literature* (Vol. 71). Oxford, UK: Blackwell.
- Feng, S., Banerjee, R., & Choi, Y. (2012). Characterizing stylistic elements in syntactic structure. In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 1522–1533). Association for Computational Linguistics.
- Finn, E. (2012). New literary cultures: Mapping the digital networks of Toni Morrison. In A. Lang (Ed.), *From Codex to Hypertext: Reading at the Turn of the Twenty-First Century* (pp.177–202). Amherst, MA: University of Massachusetts Press.
- Firth, J. (1957). *A synopsis of linguistic theory: studies in linguistic analysis*. Oxford, UK: Blackwell.
- Foucault, M. (2005). The hermeneutics of the subject: Lectures at the College de France. *Michel Foucault: Ethics, Subjectivity and Truth*. London, UK: Penguin.
- Fox, A. B., Rosen, J., & Crawford, M. (2009). Distractions, distractions: Does instant messaging affect college students' performance on a concurrent reading comprehension task? *CyberPsychology & Behavior*, 12(1), 51–53.
- Francis, W. N., & Kucera, H. (1979). Brown corpus manual. *Brown University*, 15.

- Freedman, S., Hull, G., Higgs, J., & Booten, K. (2016). Teaching writing in a digital and global age: Toward access, learning, and development for all. In D. H. Gitomer & C. A. Bell, (Eds.), *Handbook of Research on Teaching*, 5th Edition. Washington, DC: American Educational Research Association.
- Fussell, P. (1979). *Poetic meter and poetic form*. Columbus, OH: McGraw-Hill Education.
- Gaetani, G. (2015). The notable art of misquoting Camus – from its origins to the internet era. *Journal of Camus Studies*, 37-50.
- Garcia, A., Cantrill, C., Filipiak, D., Hunt, B., Lee, C., Mirra, N., ... Peppler, K.(Eds.). (2014). *Teaching in the connected learning classroom*. Irvine, CA: Digital Media and Learning Research Hub. Retrieved from <https://dmlhub.net/publications/teaching-connected-learning-classroom>
- Gawryjolek, J. J. (2009). *Automated annotation and visualization of rhetorical figures*. Waterloo, Canada: University of Waterloo Press.
- Gee, J. P. (2005). Semiotic social spaces and affinity spaces. In D. Barton & K. Tusting (Eds.), *Beyond communities of practice: Language power and social context* (pp. 214–232). Cambridge, UK: Cambridge University Press.
- Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method*. New York, NY: Routledge.
- Geertz, C. (1994). Thick description: Toward an interpretive theory of culture. *Readings in the Philosophy of Social Science*, 213–231.
- Genette, G. (1997). *Paratexts: Thresholds of interpretation* (Vol. 20). Cambridge, UK: Cambridge University Press.
- Gower, J. C., & Ross, G. (1969). Minimum spanning trees and single linkage cluster analysis. *Applied Statistics*, 54–64.
- Green, J., & Jenkins, H. (2009). The moral economy of web 2.0: Audience research and convergence culture. In J. Holt & A. Perren (Eds.), *Media industries: History, theory, and method* (pp. 213–225). West Sussex, UK: Blackwell.
- Greenwood, A., Perrin, A. & Duggan, M. (2016). *Demographics of social media users*. Pew Research Center (November 11, 2016). Retrieved from <http://www.pewinternet.org/2015/08/19/the-demographics-of-social-media-users/>
- Guillory, J. (2013). *Cultural capital: The problem of literary canon formation*. Chicago, IL: University of Chicago Press.
- Halliday, M. A. (1971). Linguistic function and literary style: An inquiry into the language of William Golding's the inheritors. In *Literary style: A symposium* (Vol. 339). New York, NY: Oxford University Press.
- Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in English*. Abingdon, UK: Routledge.
- Hamlin, A. K. (2017). Critical reading in art, art history, and wikipedia. In R DiYanni & A. Borst (Eds.), *Critical reading across the curriculum: Humanities* (pp. 104-122). Hoboken, NJ: John Wiley & Sons, Inc.
- Havelock, E. A. (1986). *The muse learns to write: Reflections on orality and literacy from antiquity to the present*. New Haven, CT: Yale University Press.
- Havens, E., & Mooney, J. E. (2001). *Commonplace books: A history of manuscripts and printed books from antiquity to the twentieth century*. Beinecke Rare Book and Manuscript Library, Yale University.
- Hayles, N. K. (2007). Hyper and deep attention: The generational divide in cognitive modes. *Profession*, 187–199.

- Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. Cambridge, UK: Cambridge University Press.
- Hellekson, K., & Busse, K. (Eds.). (2006). *Fan fiction and fan communities in the age of the internet: New essays*. Jefferson, NC: McFarland & Co, Inc.
- Hermida, A., Lewis, S. C., & Zamith, R. (2014). Sourcing the Arab Spring: A case study of Andy Carvin's sources on Twitter during the Tunisian and Egyptian revolutions. *Journal of Computer-Mediated Communication*, 19(3), 479-499.
- Hillman, S., Procyk, J., & Neustaedter, C. (2014). Tumblr fandoms, community & culture. In *Proceedings of the companion publication of the 17th ACM conference on computer supported cooperative work & social computing* (pp. 285–288). ACM.
- Hine, C. (2000). *Virtual ethnography*. London, UK: Sage Publications.
- Hoey, M. (1983). *On the surface of discourse*. London, UK: Allen & Unwin.
- Honnibal, M., & Johnson, M. (2015). An Improved Non-monotonic Transition System for Dependency Parsing. In *EMNLP* (pp. 1373-1378).
- Höpfl, H. (1983). Isms. *British Journal of Political Science*, 13(01), 1–17.
- Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 168–177). ACM.
- Hull, G. A., & Schultz, K. (Eds.). (2002). *School's out: Bridging out-of-school literacies with classroom practice* (Vol. 60). New York, NY: Teachers College Press.
- Hull, G. A., & Stornaiuolo, A. (2010). Literate arts in a global world: Reframing social networking as cosmopolitan practice. *Journal of Adolescent & Adult Literacy*, 54(2), 85–97.
- Ikeda, K., Hattori, G., Ono, C., Asoh, H., & Higashino, T. (2013). Twitter user profiling based on text and community mining for market analysis. *Knowledge-Based Systems*, 51, 35-47.
- Ito, M., Baumer, S., Bittanti, M., Cody, R., Stephenson, B. H., Horst, H. A., ... others. (2009). *Hanging out, messing around, and geeking out: Kids living and learning with new media*. Cambridge, MA: MIT Press.
- Ito, M., Gutierrez, K., Livingstone, S., Penuel, B., Rhodes, J., Salen, K., ... Watkins, S. C. (2013). *Connected learning: An agenda for research and design*. Digital Media Learning Research Hub.
- Jacobsen, W. C., & Forste, R. (2011). The wired generation: Academic and social outcomes of electronic media use among university students. *Cyberpsychology, Behavior, and Social Networking*, 14(5), 275–280.
- Jakobson, R. (1960). Linguistics and poetics. In *Style in language* (pp. 350-377). Cambridge, MA: MIT Press.
- Jenkins, H. (2012). *Textual poachers: Television fans and participatory culture*. Abingdon, UK: Routledge.
- Jenkins, H., Ford, S., & Green, J. (2013). *Spreadable media: Creating value and meaning in a networked culture*. New York, NY: NYU Press.
- Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32(3), 241–254.
- Jones, E., Oliphant, T., Peterson, P., & others. (2001–2001--). SciPy: Open source scientific tools for Python. Retrieved from <http://www.scipy.org/>.
- Junco, R. (2012). Too much face and not enough books: The relationship between multiple indices of facebook use and academic performance. *Computers in Human Behavior*, 28(1), 187–198.

- Junco, R., & Cotten, S. R. (2012). No a 4 u: The relationship between multitasking and academic performance. *Computers & Education*, 59(2), 505–514.
- Karpovich, A. (2006). The audience as editor. The role of beta readers in online fan fiction communities. In K. Hellekson & K. Busse (Eds.), *Fan fiction and fan communities in the age of the internet: New essays* (pp. 171-188). Jefferson, NC: McFarland & Co, Inc.
- King, P. M., & Kitchener, K. S. (1994). *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults*. San Francisco, CA: Jossey-Bass.
- Kirschner, P. A., & Karpinski, A. C. (2010). Facebook and academic performance. *Computers in Human Behavior*, 26(6), 1237–1245.
- Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1* (pp. 423-430). Association for Computational Linguistics.
- Knobel, M. & Lankshear, C. (2003). Researching young children's out-of-school literacy practices. In N. Hall, J. Larson, & J. Marsh (Eds.), *Handbook of early childhood literacy* (pp. 51-65). London, UK: Sage Publishing.
- Knobel, M. & Lankshear, C. (Eds.). (2007). *A new literacies sampler* (Vol. 29). New York, NY: Peter Lang.
- Knoeller, C. (1998). *Voicing ourselves: Whose words we use when we talk about books*. Albany, NY: SUNY Press.
- Kraushaar, J. M., & Novak, D. C. (2010). Examining the effects of student multitasking with laptops during the lecture. *Journal of Information Systems Education*, 21(2), 241.
- Kress, G. R., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. East Sussex, UK: Psychology Press.
- Lakkaraju, H., McAuley, J. J., & Leskovec, J. (2013). What's in a name? Understanding the interplay between titles, content, and communities in social media. *ICWSM*, 1(2), 3.
- Lam, W. S. E. (2000). L2 literacy and the design of the self: A case study of a teenager writing on the internet. *Tesol Quarterly*, 34(3), 457–482.
- Landow, G. P. (1991). *HyperText: The convergence of contemporary critical theory and technology (parallax: Re-visions of culture and society series)*. Baltimore, MD: Johns Hopkins University Press.
- Lankshear, C., & Knobel, M. (2007). Sampling "the new" in new literacies. In *A New Literacies Sampler* (pp. 1–24). New York, NY: Peter Lang.
- Lankshear, C., & Knobel, M. (2011). *New literacies: Everyday practices and social learning*. Berkshire, UK: McGraw-Hill Education.
- Levine, L. E., Waite, B. M., & Bowman, L. L. (2007). Electronic media use, reading, and academic distractibility in college youth. *CyberPsychology & Behavior*, 10(4), 560–566.
- Lin, L., Robertson, T., & Lee, J. (2009). Reading performances between novices and experts in different media multitasking environments. *Computers in the Schools*, 26(3), 169–186.
- Mangen, A., Walgermo, B. R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61–68.
- Manikonda, L., Hu, Y., & Kambhampati, S. (2014). Analyzing user activities, demographics, social network structure and user-generated content on Instagram. *arXiv preprint arXiv:1410.8099*.
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional



- theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse*, 8(3), 243–281.
- Marx, K. & Engels, F. (1942). *Selected correspondence, 1846-1896* (D. Torr, trans.). New York: International Publishers.
- Marwick, A., & Boyd, D. (2011). To see and be seen: Celebrity practice on Twitter. *Convergence*, 17(2), 139-158.
- McEnery, A. M., & Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh, UK: Edinburgh University Press.
- Mitchell, J. C. (1984). Typicality and the case study. *Ethnographic Research: A Guide to General Conduct*, 238241.
- Moss, A. (1996). *Printed commonplace-books and the structuring of renaissance thought*. New York, NY: Oxford University Press.
- Murray, D. C. (2015). Notes to self: the visual culture of selfies in the age of social media. *Consumption Markets & Culture*, 18(6), 490-516.
- Murray, S. (2015). Charting the digital literary sphere. *Contemporary Literature*, 56(2), 311–339.
- Myers, G. (2016). Everyday oracles: Authors on twitter. *Celebrity Studies*, 7(4), 476–492.
- Naik, Y. (2012). Finding good reads on goodreads: Readers take rA into their own hands. *Reference & User Services Quarterly*, 51(4), 319.
- Nakamura, L. (2013). "Words with friends": Socially networked reading on goodreads. *PMLA*, 128(1), 238–243.
- Nelson, T. H. (1993). *Literary machines 93.1*. Sausalito, CA: Mindful Press.
- Ong, W. J. (2013). *Orality and literacy*. Abingdon, UK: Routledge. (Original work published 1982)
- Page, R. (2013). Seriality and storytelling in social media. *StoryWorlds: A Journal of Narrative Studies*, 5(1), 31–54.
- Perrin, A. (2016). *Book reading 2016*. Pew Research Center. Retrieved from <http://www.pewinternet.org/2016/09/01/book-readubg-2016/>
- Peterson, R. A. (1992). Understanding audience segmentation: From elite and mass to omnivore and univore. *Poetics*, 21(4), 243–258.
- Peterson, R. A., & Kern, R. M. (1996). Changing highbrow taste: From snob to omnivore. *American Sociological Review*, 900–907.
- Pew Research Center (2015). Demographics of Tumblr. Retrieved from [http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19\\_social-media-update\\_05/](http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_05/).
- Pratt, M. L. (1977). *Toward a speech act theory of literary discourse*. Bloomington, IN: Indiana University Press.
- Phillips, F. (2011). The Star Wars franchise, fan edits, and lucasfilm. *Transformative Works and Cultures*, 9.
- Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. Urbana, IL: University of Illinois Press.
- Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *Proceedings of the workshop on comparing corpora* (pp. 1–6). Association for Computational Linguistics.
- Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Generation M: Media in the lives of 8-to 18-year-olds*. Henry J. Kaiser Family Foundation (Jan 20, 2010). Retrieved from

- <http://www.kff.org/other/report/generation-m2-media-in-the-lives-of-8-to-18-year-olds/>
- Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. *Computers in Human Behavior*, 29(3), 948–958.
- Rowberry, S. P. (2016). Commonplacing the public domain: Reading the classics socially on the kindle. *Language and Literature*, 25(3), 211–225.
- Saxton, G. D., Niyirora, J. N., Guo, C., & Waters, R. D. (2015). Advocating for change: The strategic use of hashtags in social media advocacy. *Advances in Social Work*, 16(1), 154–169.
- Saldaña, J. (2015). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage Publishing.
- Schmitt, C. B., Skinner, Q., & Kessler, E. (1988). *The cambridge history of renaissance philosophy*. Cambridge, UK: Cambridge University Press.
- Scholz, T. (2008). Market ideology and the myths of web 2.0. *First Monday*, 13(3).
- Scolari, C. A. (2009). Transmedia storytelling: Implicit consumers, narrative worlds, and branding in contemporary media production. *International Journal of Communication*, 3, 586–606. <http://www.ijoc.org/index.php/ijoc/article/view/477>
- Scribner, S., Cole, M., & Cole, M. (1981). *The psychology of literacy* (Vol. 198). Cambridge, MA: Harvard University Press.
- Segal, E. M., & others. (1995). Narrative comprehension and the role of deictic shift theory. *Deixis in Narrative: A Cognitive Science Perspective*, 3–17.
- Senft, T. M. (2013). Microcelebrity and the branded self. *A Companion to New Media Dynamics*, 346–354.
- Shifman, L. (2013). *Memes in digital culture*. Cambridge, MA: MIT Press.
- Shor, I. (1999). What is critical literacy. *Journal for Pedagogy, Pluralism & Practice*, 4(1), 1–26.
- Shorey, S. (2015). *Fragmentary girls: Selective expression on the Tumblr platform* (Master's thesis). University of Massachusetts, Amherst.
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778.
- Spradley, J. P. (2016). *The ethnographic interview*. Long Grove, IL: Waveland Press.
- Stein, L. E. (2006). "This dratted thing": Fannish storytelling through new media. In K. Hellekson & K. Busse (Eds.), *Fan Fiction and Fan Communities in the Age of the Internet* (pp. 245–260). Jefferson, NC: McFarland & Co, Inc..
- Sterponi, L. (2007). Clandestine interactional reading: Intertextuality and double-voicing under the desk. *Linguistics and Education*, 18(1), 1–23.
- Street, B. (2001). Literacy events and literacy practices: theory and practice in the New Literacy Studies. In M. Martin-Jones & K. Jones (Eds.), *Multilingual literacies: Reading and writing different worlds* (pp. 17–30). Amsterdam, The Netherlands: John Benjamins Publishing.
- Striphas, T. (2009). *The late age of print: Everyday book culture from consumerism to control*. New York, NY: Columbia University Press.
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford, UK: Blackwell.
- Stubbs, M. (2001). Computer-assisted text and corpus analysis: lexical cohesion and communicative competence. In D. Schiffrin, D. Tannen, & H.E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 304–320). Malden, MA: Blackwell.

- Tagg, C. (2012). *Discourse of text messaging: Analysis of SMS communication*. London, UK: A&C Black.
- Traugott, E. (1973). Generative semantics and the concept of literary discourse. *Journal of literary semantics*, 2(1), 5-22.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1–13.
- Thelwall, M. (2017). Reader and author gender and genre in goodreads. *Journal of Librarianship & Information Science*.
- Thelwall, M., & Kousha, K. (2017). Goodreads: A social network site for book readers. *Journal of the Association for Information Science and Technology*, 68(4), 972–983.
- Thorburn, D., & Jenkins, H. (2003). *Rethinking media change: The aesthetics of transition*. Cambridge, MA: MIT Press.
- Todorov, T. (1981). *Introduction to poetics* (Vol. 1). Minneapolis, MN: U of Minnesota Press.
- Vlieghe, J., Vandermeersche, G., & Soetaert, R. (2016). Social media in literacy education: Exploring social reading with pre-service teachers. *New Media & Society*, 18(5), 800-816.
- Vlieghe, J., Muls, J., & Rutten, K. (2016). Everybody reads: Reader engagement with literature in social media environments. *Poetics*, 54, 25–37.
- Vlieghe, J., Page, K. L., & Rutten, K. (2016). "Twitter, the most brilliant tough love editor you'll ever have." reading and writing socially during the twitter fiction festival. *First Monday*, 21(4).
- Voigt, E. B. (2009). *The art of syntax: Rhythm of thought, rhythm of song*. Minneapolis, MN: Graywolf Press.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (2012). *Thought and language*. A. Kozulin (Ed.). Cambridge, MA: MIT Press.
- Widdowson, H. G. (1992). *Practical stylistics: An approach to poetry*. New York, NY: Oxford University Press.
- Williams, R. (2014). *Keywords: A vocabulary of culture and society*. New York, NY: Oxford University Press.
- Wodak, R., & Meyer, M. (2009). *Methods for critical discourse analysis*. London, UK: Sage Publishing.
- Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers & Education*, 58(1), 365–374.
- Zhou, X. (2009). The political blogosphere in China: A content analysis of the blogs regarding the dismissal of Shanghai leader Chen Liangyu. *New Media & Society*, 11(6), 1003–1022.

## APPENDIX A QUESTIONNAIRE

*Note: Questions were open-ended unless otherwise specified.*

1.

On a scale from 0 (terrible, unquotable) to 10 (awesome, very quotable), how would you rate the following quote:

"Definitions belong to the definers, not the defined." - Toni Morrison, *Beloved*

Or:

"The more uncertainty, the more pain." - Fyodor Dostoevsky, *Notes from Underground*

[Quoters were randomly shown a version of a quote with or without authorial attribution.]

2.

What are some of the most common ways that you go about finding quotes to post?

3.

Why do you think people spend so much time and energy sharing and collecting quotes on Tumblr? What value do people get out of them?

4.

What are the qualities of a good quotation?

5.

Without looking it up, try to remember as much as you can from a quote that you find especially powerful (and, if possible, who said it or wrote it):

6.

Why do you think this quote sticks with you? When do you think about it? (Skip if you don't remember a quote.)

7.

When you see a quote on another blog from a book you haven't read, do you ever find the book to read some of it?

[Never, Rarely, Sometimes, Frequently, All the time]

8.

[If "Never" not selected in above.]

Oh yeah? What are some of the books you have discovered through quotes on Tumblr?

9.

Finally, some basic info. How would you describe your gender?

[This question was left open-ended to allow for non-binary gender identifications.]

10.

Your age?

[Participants moved a marker on a numerical scale.]

## APPENDIX B KEY CODES

### *Questionnaire*

- Question 2
  - BOOKS
  - NON-BOOK
    - OTHER MEDIA
      - e.g. TV, movies
    - SOCIAL MEDIA
      - e.g. Goodreads
- Question 3
  - SELF-EXPRESSION
  - SELF/OTHER CARE
    - FEELING CONNECTED/UNALONE (related to SELF-EXPRESSION)
    - IMPROVEMENT
      - e.g. inspiring, motivating, "eye-opening"
  - BOOKISH REASONS
    - e.g. finding, recommending, remembering books, finding other readers
- Question 4
  - FORM
    - POWERFUL LANGUAGE
      - "poetic wording"
    - BREVITY
      - "less than three sentences," "pithy"
  - MEANING
    - COGNITIVE/AFFECTIVE EFFECTS
      - INSPIRING
        - "inspiring"
      - OTHERS
        - "makes you think," "evokes strong feeling"
    - RESONANT
      - "resonates with you," "relatable to my own life"
    - UNIVERSAL
      - "[s]omething everyone can relate to in some way"
- Question 6
  - SELF-CARE
    - APPLICATION TO ONE'S OWN LIFE/STRUGGLES

### *Follow-Up Interviews*

- Question about How Quoters
  - PERSONAL INFORMATION MANAGEMENT
  - QUOTING WHILE READING versus QUOTING AFTER READING

### *BQ Photo Corpus*

- IMAGES FROM OTHER MEDIA (e.g. television, film)
- PHOTOGRAPHS OF BOOKS

- READERLY PHOTOGRAPHS
  - i.e. pictures of highlighted books
- AESTHETICIZED PHOTOGRAPHS
  - i.e. pictures of books artfully posed with accouterments of reading (e.g. coffee) or other objects (e.g. flowers)