

Copyright
by
Cindy Kyuah Chung
2009

**The Dissertation Committee for Cindy Kyuah Chung Certifies that this is the
approved version of the following dissertation:**

**PREDICTING WEIGHT LOSS IN BLOGS
USING COMPUTERIZED TEXT ANALYSIS**

Committee:

James W. Pennebaker, Supervisor

Christopher G. Beevers

Joydeep Ghosh

Samuel D. Gosling

Robert A. Josephs

**PREDICTING WEIGHT LOSS IN BLOGS
USING COMPUTERIZED TEXT ANALYSIS**

by

Cindy Kyuah Chung, B.Sc.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2009

Dedication

정경문

Acknowledgements

Heartfelt thanks to James W. Pennebaker for always supplying ammo to the six-shooters I never knew I had. In research and in life, you taught me to cowboy up, and that when I do, it'll be okay. You have led by example, and it has been my greatest honour and joy to work with you.

Thanks to the dissertation committee for shaping the project and their support throughout: Samuel D. Gosling, Christopher G. Beavers, Joydeep Ghosh, and Robert A. Josephs. Thanks to Jerry Parlee for showing me that discoveries can come from simply organizing data. Thanks also to Clinton Jones and Alexander Liu for taking a first scratch at the data with me. Many thanks to the research assistants who provided a closer look at the data: Gizem Engur, Taylor Lightfoot, Paola Rodriguez, Dianne Segovia, Vera J. Vine, and Alessandra Ytuarte. Thanks to Marlone Henderson for suggestions for future research.

Ineffable gratitude to my sister, Lois S. Chung, and my brother, A. Jinu Chung who have kept it together for me, and who made immeasurable sacrifices for me to do what I love. Thanks to 정경문, 김종금, Andrew Rhee, and Naomi Rhee for giving me the support that made this degree possible, and to Amy Taimour Ang, Sharon Choi, Yousra Qureshi, and Theodore Kim for supporting me along the way.

Thanks to my friends, collaborators, and mentors in graduate school, Ewa Kacwicz, Nairan Ramirez-Esparza, Kate G. Niederhoffer, Richard B. Slatcher, Matthias R. Mehl, Navid Hayeri, and Markus Wolf, the best minds and friends to make this experience the best it could possibly be. Special memories and thanks to Simine Vazire and Pranjal Mehta for the Paper a Month Club, and to Jenna L. Baddeley for Paper

Challenges 2009 & Quiet Hours. Thanks to those who mentored me as their own, Arthur C. Graesser, Jeffrey T. Hancock, and David I. Beaver. Good things are meaningful only if you have good people to share them with, so to Hani B. Freeman, Adrian Garcia, Katie & Jim McClarty, Sarah & Jorge Angulo, Ania & Robbie Kacewicz, Carl Spendiff, Derek Springer, Gail & Lucy & Lotte Niederhoffer, Lenny R. Vartanian, Chris Mazzucco, Yla Tausczik, Lindsay T. Graham, Brencho Hughes, Christine & Bret Chang-Schneider, Conor D. Seyle, Amit Marwaha, Christopher S. Henry, and all those who have been a part of the Pennebaker lab and the Social Personality area, sincere thanks.

PREDICTING WEIGHT LOSS USING COMPUTERIZED TEXT ANALYSIS

Publication No. _____

Cindy Kyuah Chung, Ph.D.

The University of Texas at Austin, 2009

Supervisor: James W. Pennebaker

An increasing number of people are turning to online blogging communities devoted to self-change for smoking, shopping, and other behaviors. To understand processes underlying effective self-change, the current project tracked the language and social dynamics of a dieting blog community using computerized text analysis. Three research questions were asked: What predicts weight loss in blogs? What changes in blogging predict weight loss? Can we predict dropping out or successful weight loss based on the first two entries? A community of blogs devoted to weight loss was examined (n = 2530). Most bloggers were female, and on average, around 30 years old, approximately 200 pounds, with a goal weight of about 140 pounds. A sample of blogs by females that had blogged at least 15 entries within the first 15 weeks of blogging resulted in a total of 186 blogs, representing over 9,200 entries for analysis.

Computerized text analysis was used to examine language for rates of self-focus, emotionality, cognitive processing, keeping food diaries, and social support. Rates of

blogging were assessed by word counts, number of active weeks, and mean entries per week. Social support was assessed through the use of social words, the size of the social network, along with the positivity and negativity of the comments. The discrepancy between start and goal weight was also assessed. The results suggested that having larger weight loss goals and blogging about personal events was a more effective weight loss strategy than keeping an online food intake diary. The degree to which bloggers were socially integrated with the blog community was found to be a potent predictor of weight loss. Online components of behavioral treatment programs could encourage dieters to browse and comment on other dieters' progress, and to share personal narratives rather than simply focusing on the benefits of food intake diaries, nutrition, and exercise. The current project points to the power of computerized text analytic tools to address important theoretical and practical social psychological issues that are evolving on the internet. Specifically, language analysis methods can identify which dimensions of blogging communities can help or hinder self-change processes.

Table of Contents

List of Tables	XII
List of Figures	XIV
Chapter 1 Introduction	1
Previous Work: Self-regulation in Micronarratives	3
Measurement by Word Counts	5
Word Use and its Links to Psychological State	10
Summary	18
Chapter 2 The Present Study: Predicting Weight Loss Based on Natural Language in Blogs	20
Self-Focus: First Person Singular Pronouns	21
Emotion Word Use	22
Cognitive Word Use	23
Self-Monitoring	24
Social Support.....	26
Summary	28
Research question 1: What Predicts Weight Loss in Blogs?	29
Research question 2: What Changes in Blogging Predict Weight Loss?	31
Research question 2: Can We Predict Dropping Out or Weight Loss from the First Two Entries.....	32
Chapter 3 The Weight Loss Blogs Corpus and Sample.....	34
Weight Loss Blogs Corpus	34
Sample	44
Target Measures	48
Summary	55
Chapter 4 Research Question 1: What Predicts Weight Loss in Blogs?.....	57
Chapter 5 Research Question 2: What Changes in Blogging Predict Weight Loss in Blogs?	65

Chapter 6 Research Question 3: Can We Predict Dropping Out or Weight Loss from the First Two Entries?	74
Dropouts vs. Dedicated Bloggers	74
Weight Loss by Dedicated Bloggers	82
Chapter 7 General Discussion.....	89
Appendices.....	109
Appendix A: Instructions to Judges for Ratings of Entries	109
Appendix B: Correlations of Judges’ Ratings of Food Diaries with LIWC’s Biological Mechanisms Category and Subcategories.....	111
Appendix C: Correlations of Judges’ Ratings of Social Support with LIWC’s Social Category and Subcategories.....	112
Appendix D: Correlations of Judges’ Ratings of Making Realizations with LIWC’s Cognitive Mechanism Subcategories.....	113
Appendix E: Correlations of Judges’ Ratings of Expressing and Suppressing Negative and Positive Emotions with LIWC’s Affect Categories and Subcategories	114
Appendix F: Correlations of Judges’ Ratings of Self-Focus with Pronoun Use	115
Appendix G: Correlations between Judges’ Ratings of Items	116
Appendix H: Correlations of Judges’ Ratings with all LIWC Categories..	117
Appendix I: Correlations of Percent Body Weight Loss over the 15 Weeks with all LIWC Categories, other Blogging Rates, and Comment Counts	119
Appendix J: Correlations between Target Variables in the First 15 Weeks	123
Appendix K: Original Regression for Research Question 1 with All Cases (i.e. including Outliers)	124
Appendix L: Differences Scores of all LIWC Categories, other Blogging Rates, and Comment Counts over the First and Second Half of Blogging.....	125
Appendix M: Correlations of Percent Body Weight Loss Over the 15 Weeks with Differences Scores of all LIWC Categories, other Blogging Rates, and Comment Counts over the First and Second Half of Blogging.....	129

Appendix N: Correlations between Difference Scores across Halves of Blogging for the Target Variables	133
Appendix O: Original Regression for Research Question 2 with All Cases (i.e. including Outliers)	134
Appendix P: Mean Differences between Dropouts and Dedicated Bloggers on all LIWC Categories for the First Two Entries.....	135
Appendix Q: Correlations between Target Variables within the First Two Entries by Dropouts	138
Appendix R: Correlations between Target Variables within the First Two Entries by Dedicated Bloggers.....	139
Appendix S: Correlations of Target Variables during the First Two Entries with Group Membership (Dropouts = 0; Dedicated Bloggers = 1) ...	140
Appendix T: Correlations of Percent Body Weight Loss over the 15 Weeks with all LIWC Categories, other Blogging Rates, and Comment Counts Measured within the First Two Entries of Dedicated Bloggers ..	141
Appendix U: Original Regression for Research Question 3 (Predicting Weight Loss in the Dedicated Bloggers based on the First Two Entries) with All Cases (i.e. including Outliers).....	144
References	145
Vita	157

List of Tables

Table 1: Descriptive Information on www.dietdiaries.com Corpus	35
Table 2: Number of Blogs with 1 to 15 Entries	37
Table 3: Number of Blogs with 16 Entries or More	38
Table 4: Exclusion Criteria for Sample	46
Table 5: Week of Final Post During the 15 Week Sample Period for Blogs in the Sample	47
Table 6: Descriptive Information on www.dietdiaries.com Sample (n = 186)	48
Table 7: Inter-Rater Reliabilities of Judge’s Ratings of Entries and Correlations to LIWC Categories	51
Table 8: Correlations between Percent Body Weight Loss and Mean Levels of Target Variables During the 15 Week Period.....	58
Table 9: Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss	61
Table 10: Mean Differences on Target Variables in the First Half and the Second Half of Blogging	66
Table 11: Correlations between Percent Body Weight Loss and Difference Scores between First and Second Halves of Blogging During 15 Weeks of Blogging	70
Table 12: Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight LOSS	71
Table 13: Mean Age and Weight Differences between Dropouts and Dedicated Bloggers	76

Table 14: Mean Differences between Dropouts and Dedicated Bloggers on Target Variables for the First Two Entries	77
Table 15: Beta Weights and their Statistical Significance in a Regression Equation Predicting Group Membership (i.e. Dropout or Dedicated Blogger)	79
Table 16: Classification Table for Regression Model Predicting Dropout or Dedication....	80
Table 18: Correlations between Percent Body Weight Loss and Target Variables Within the First Two Entries	83
Table 19: Beta Weights and their Statistical Significance in a Regression Equation Predicting Weight Loss	85

List of Figures

Figure 1: President George W. Bush’s Use of First Person Singular Pronouns During his Months in Office.	15
Figure 2: Histogram of Age at Start of Blog for Bloggers in the Corpus where Starting Age Information was Available (n = 1331).....	37
Figure 3: Histogram of Total Number of Posts by Bloggers in the Corpus Who Blogged One Week or Less (51.58% of all Bloggers).....	39
Figure 4: Histogram for Blogs With Durations Greater than One Week	39
Figure 5: Histogram of Goal Weights in Pounds for those Blogs in which Goal Weight Information was Available (n = 1675)	41
Figure 6: Histogram of Starting Weights for Bloggers (i.e. Weight at First Entry; n = 1325).....	42
Figure 7: Scatterplot of Starting Weights and Goal Weights	43
Figure 8: Scatterplot of Weight Loss for Bloggers in the Corpus for whom Start Weight and End Weight Information were available.....	43

Chapter 1: Introduction

The use of blogs for monitoring and documenting the process of self-change is an emerging phenomenon (e.g. Harmon, 2003; Leland, 2007; National Public Radio, 2007; Williams, 2006; Zapata, 2007). Thousands of bloggers are chronicling their daily successes and struggles with goals related to weight loss. With each entry, information on weight lost or gained is documented. Anyone can start an online journal about their own weight loss, read and respond to other blogs on weight loss, and keep a daily record of their weight. Some people use the blogs as food intake diaries, some use the blogs as outlets to vent, to get support, to get advice, to share their success stories, to reflect on daily events, or for all of the above reasons.

The archival and public nature of blogs makes them an ecologically valid source for collecting and examining naturalistic behaviors over time. Blogs are often centered on a particular topic (e.g. politics, hobbies, celebrities) or a particular goal (e.g. weight loss, debt removal, marathon training). Oftentimes, blogs on the same topic or goal are linked to one another making up a blog community. Therefore, it is possible to track how large numbers of people are thinking about any given topic, responding to world events, or achieving a goal by examining word use within blogs devoted to a particular blog community.

Although there have been clinical trials on the efficacy of self-help and commercial weight loss programs (e.g. Heshka, Anderson, Atkinson, Greenway, Hill, Phinney, et al., 2003), experimental research on overeating in the laboratory (see Polivy & Herman, 1985) and examinations of daily intake diaries (e.g. de Castro, 1994; Legg,

Puri, & Thomas, 2000), there has been little research on the factors that lead to successful weight loss in everyday life. Self-change endeavors such as weight loss are likely to be salient throughout the day. Accordingly, the process of self-change needs to be examined more closely for what leads to such gross failure rates, and equally important, how people succeed.

With the extant archives of natural language samples on the Internet, it is possible to access a large group of dieters who track their daily thoughts and feelings over time. This dissertation will track language markers of successful self-change in blogs devoted to weight loss. The online weight loss community allows us to track the degree to which dieters communicate with each other, providing markers of some basic social processes that may influence both language and dieting. The study of blogs devoted to weight loss offers a unique opportunity to examine the natural course of known successful strategies in self-regulation, along with an exploratory look at any new language patterns that might characterize the self-change process.

To understand the process of self-change, a weight loss blog community was examined using computerized text analysis and an analysis of the social network. This dissertation begins with a brief review on what has been learned about psychological states using computerized word count approaches on natural language samples. A handful of studies that have begun to track language use over time are described. Finally, a series of research questions were posed to predict weight loss in a community of diet blogs. Thus, this dissertation showcases how information within blog communities can be measured with the aid of computerized text analysis.

PREVIOUS WORK: SELF-CHANGE IN MICRONARRATIVES

What might be learned if people were to write about their self-change goals? In order to learn more about the experience of self-change without formal or clinical treatment, Heatherton and Nichols (1994) randomly assigned one group of participants to write a micronarrative (i.e. an autobiographical account of a life event) about an actual time in their lives when they experienced a sudden and dramatic life change. The other participants were asked to describe what prevented them from achieving a life change. Both conditions were given examples of life changes such as quitting smoking, quitting a job, starting a new career, or leaving an abusive relationship.

Judges' ratings showed that successful change stories were more likely to mention actively seeking out help, and were more likely to report receiving help and social support. Although the two groups did not differ in the degree to which other people commented on their life change, there were more instances of others commenting on the process of change for successful change attempts. Successful change stories were more likely to mention seeking out self-help or educational materials, making a public declaration that one would change, and changing the immediate environment. In brief, social support was highly salient to successful changers.

Sobell and colleagues (2001) conducted a computer-assisted content analysis of open-ended interviews of self-change across cultures (Canada and Switzerland) and across domains (alcohol, cocaine, and heroin). Using a software program, Textpack (Mohler & Zull, 1998), the researchers developed the Natural Recovery Dictionary (NRD), a dictionary made up of key terms related to recovery from drugs and alcohol. By

counting the frequency of key terms in various content categories for each interview, the researchers found evidence for linguistic markers of spontaneous recovery (i.e. without formal treatment) from substance abuse across cultures. Words falling into their categories of cognitive evaluations (e.g. belief, commit, consider, realize) and behavioral monitoring (e.g. attempt, control, resolve, try) were mentioned most frequently.

Using both qualitative and the NRD computerized text analysis of smokers' self-change micronarratives, Helvig and his colleagues (2006) found that quitters who had maintained their non-smoking status used significantly more affective statements than did failed quitters. In addition, maintained quitters used significantly more statements related to pro-smoking cessation social support than did failed quitters.

The above studies on narratives of self-change found that emotions, monitoring behavior, and social support were important determinants in the self-change process. However, there are a few limitations with the micronarrative approach to studying self-change. Qualitative studies have not been systematic in their coding categories, leaving the field with multiple classifications or narrative excerpts that cannot be easily compared across studies or cultures (Sobell, et al., 2001). The NRD contained very few words, accounting for very few of all words in the interviews. The measurement of words per category (i.e. by absolute frequency) did not account for the overall word count in the interviews. Percentages of total word count would have been more representative of the uses of word categories within the interviews than word frequencies due to varying interview lengths. In addition, no other known studies have used the NRD, making the results difficult to compare across other studies. Lastly, as in most examinations of

reasons for self-change, the study used a retrospective self-report method. Although it is important to note that the participants' perceptions of their reasons for change were likely highly influential in the maintenance of change and avoidance of relapse (Heatherton and Nichols, 1994; Helvig et al., 2001), it is possible that memory biases and distortions may have systematically arose in response to recall of success or failure narratives. Therefore, while the micronarrative approach is promising for identifying possible ways that people recall self-change attempts, new sources that document the process of self-change online are available.

MEASUREMENT BY WORD COUNTS

Recent advancements in computerized text analyses have allowed psychologists to develop new ways of thinking about how language is associated with various thoughts and emotions (Biber, 1988; Pennebaker, Mehl, & Niederhoffer, 2003). Unlike qualitative text analyses involving manual coding of themes by human judges as in the micronarrative analysis, computerized text analytic tools have enabled more efficient, more reliable, and novel ways of analyzing larger volumes of text (Mehl, 2006a). Although there are complex algorithmic approaches to text analysis using word patterns, this dissertation will focus on what can be learned about psychological processes through a broad and increasingly popular approach to quantitative text analysis using simple word counts.

As early as the 1960s, crude word counting methods were devised that attempted to capture people's basic motives and emotions (e.g., Stone, Dunphy, Smith, & Ogilvie, 1966). Weintraub (1981) hand counted thousands of documents and speech samples to

provide compelling evidence that people's use of certain parts of speech were diagnostic of their emotional well-being. Similar research in the measurement of natural language use in psychology followed (Martindale, 1990; Mergenthaler, 1996; Gottschalk, Winget, & Glaser, 1969). Today, there is an increasing number of applications of word analyses in clinical (e.g. Gottschalk, 2000; Ramirez-Esparza, Chung, Kacewicz, & Pennebaker, 2008), forensic (e.g. Adams, 2002; 2004; Pennebaker & Huddle, unpublished data), cultural (e.g. Boroditsky, 2001; Tsai, Simenova, & Watanabe, 2004) and personality assessments (e.g. Mehl, Gosling, & Pennebaker, 2006; Oberlander & Gill, 2004; Pennebaker & King, 1999).

Word count approaches typically rely on a set of dictionaries with pre-categorized terms. The categories can be grammatical categories (e.g., adverbs, nouns, prepositions, verbs, etc.) or content categories (e.g. positive emotions, cognitive words, social words, etc.). While grammatical categories are fixed (i.e. entries belong in one or multiple known categories), content categories are formed by judges' votes on whether or not each word belongs in a category. Computerized software can then be programmed to categorize words appearing in text according to the dictionary that it references. Accordingly, these programs typically allow for the use of new, user-defined dictionaries, enabling broader or more specific sampling of word categories. For example, as described in the previous section, the National Recovery Dictionary was used with Textpack to count the occurrences of cognitive evaluation words (e.g. belief, cognizant, recognize, think), support words (e.g., ex-alcoholic, marriage, relationship, ultimatum), along with symptom (e.g. cirrhosis, hallucinations, tremors, withdrawals) and alcohol

related words (e.g. alcoholic, bar, booze, sober) within self-change narratives (Sobell et al., 2001).

Among grammatical categories, word counts have recently focused on function words or closed-class words (Chung & Pennebaker, 2007). Function words include pronouns, prepositions, articles, conjunctions, and auxiliary verbs, and account for over half of the words we use in daily speech (Rochon, Saffran, Berndt, & Schwartz, 2000). Function words are generally very short (usually 1-4 letters), are spoken quickly (at a speed of 100-300 milliseconds – the rate often used in lab studies testing priming or subliminal perception) - and read quickly (Van Petten & Kutas, 1991). Given that function words are so difficult to control, examining the use of these words in natural language samples has provided a non-reactive way to explore social and personality processes.

Among content categories, which are largely made up of nouns, regular verbs, adverbs, and adjectives (also known as open-class words), the most commonly studied category has been emotion words. While the taxonomy of emotions is debatable, a commonly accepted distinction in many theories of emotions is in the valence of emotions (Fontaine, Scherer, Roesch, & Ellsworth, 2007). That is, people generally distinguish between positive and negative emotions, and this distinction appears in most word count tools developed by psychologists (e.g. Stone et al., 1966; Pennebaker, Booth, & Francis, 2007). Other content categories are associated with social words (e.g. communication, family, friends) which tap the degree to which people are attending to their social worlds. Cognitive mechanism categories (e.g. causal word categories, insight

categories, tentativeness) tap the degree to which people are thinking about causes, consequences, or conflicts about a particular topic. Note that certain function words are also grouped with cognitive mechanism word categories. For example, LIWC's discrepancy category includes modal verbs such as should, ought, and must, along with content words such as hope, mistake and regret.

An increasingly popular tool used for text analysis within psychology is Linguistic Inquiry and Word Count (LIWC2007; Pennebaker, et al., 2007). LIWC is a computerized word counting tool that searches for approximately 4,000 words and word stems, and categorizes them into grammatical (e.g. articles, numbers, pronouns), or content (e.g. achievement, cognitive, death, emotions, social) categories. Results are reported as a percentage of words in a given text file, indicating the degree to which a particular category was used. The words in LIWC categories have previously been validated by independent judges, and the use of the categories within texts have been shown to be reliable markers for a number of psychologically meaningful constructs (Pennebaker, et al., 2003).

For example, using LIWC, word counts have been shown to have modest yet reliable links to personality and demographics (Pennebaker & King, 1999). Newman and colleagues (2008) found that across 14,000 different genres of texts, women tend to use more personal pronouns and social words, whereas men tend to use more articles, numbers, and fewer verbs. Together, these suggest that women are more socially oriented, and that men tend to focus more on objects. The relationships between word use and various psychological states using LIWC have been found in spoken language (Mehl

et al., 2006), in published literature (Pennebaker & Stone, 2003), and in computer-mediated communication (Oberlander & Gill, 2002). There is also evidence that word counts are diagnostic of various psychiatric disorders, and can reflect specific psychotic symptoms (Junghaenel, Smyth, & Santner, 2008; Oxman, Rosenberg, Schnurr, & Tucker, 1988). For example, Junghaenel et al (2008) found that psychotic patients tend to use fewer cognitive mechanism words and communication words than do people who are not mentally ill, reflecting psychotic patients' tendencies to avoid in-depth processing and their general disconnect from social bonds.

Larger correlations have been found for word use and specific, directly observable behaviors more than with self-ratings on broad personality traits (Fast & Funder, 2008). For example, Fast and Funder (2008) found that the use of LIWC's certainty category was highly correlated with behavior ratings of speaking fluently and loudly. The use of LIWC's sexual category was highly correlated with behavior ratings of being talkative and dominant, which reflected a high need for attention. Overall, these studies suggest that word use is reflective of thoughts and behaviors that characterize psychological states; word counts provide meaningful measures for a variety of thoughts and behaviors.

While previous work was focused on establishing reliable individual differences in word use, the corpora used were relatively small and the work exploratory. With the growing number of genres that archive text on the Internet, the corpora for psychological analyses have become larger and more complex. Accordingly, research in the psychological analysis of natural language texts has been extended to interactions with more speakers with larger audiences, across languages, cultures, and time.

WORD USE AND ITS LINKS TO PSYCHOLOGICAL STATES

Beyond examining differences between in language use as a function of demographics or personality traits, research using computerized text analysis tools has been extended to tracking behaviors over time.

Tracking Emotional States over Time

The Internet offers a naturalistic source for tracking the thoughts and feelings of millions of people. A study of blogs written before and after 9/11 demonstrates the utility of examining natural language use in blogs. Cohn, Mehl, and Pennebaker (2001) downloaded over 70,000 blogs from livejournal.com, a site for keepers of public online journals. For each of the approximate 1,100 bloggers, the researchers downloaded blog entries for 2 months prior to and after 9/11 to examine baseline behavior, as well as immediate and long term reactions to the event. Using LIWC, they tracked the use of various word categories including emotion words, cognitive mechanism words, and words referring to others. Word usage followed a social coping model (Pennebaker & Harber, 1993), wherein the words people used reflected an increase in positive emotion words (e.g. grateful, happy, love, sharing) , and references to others following the event (e.g. she, they, family, friends), without fully returning to baseline levels within 2 months of the event. An elevated use of negative emotion words (e.g. depressed, hate, revenge, worry) after 9/11 returned to baseline within a few days. Taken together, these suggested that the terrorist attacks had the effect of bringing people together, and that people were more focused on positive than negative experiences. This pattern of effects was also

found in American Online (AOL) real-time chat room transcripts in the four weeks following the death of Princess Diana (Stone & Pennebaker, 2002).

These studies found evidence for the effects of traumatic events on thoughts and feelings within an individual across time, and across a large number of individuals. Through blog and chat forums, which are updated frequently during the period over which the events had taken place, the researchers bypassed some of the memory biases and distortions that arise in recall situations on questionnaires. In addition, examining language use in blogs presented a non-reactive and unobtrusive way to bypass some of the response rate and selection biases that result when online users are invited to participate in a research survey.

Although there were some limitations to their data from an online sample, such as the self-selection and anonymity of participants, lack of verifiable personal information, and an uneven number of contributions by each participant, these were offset by the ecological validity and richness of their corpus. By examining real world natural language as it occurs on the Internet, the authors were able to track the lifespan of a topic, tap the virtual social world of thousands of participants, and provide validation for the social stage coping model.

Tracking Self and Social Focus Over Time

Word use has shown to be reflective of underlying psychological, and even physiological changes. For example, one study examined between 1 to 2 years of personal journals and outgoing emails of two adults (one biological male and one biological female) who were undergoing testosterone therapy for different reasons

(Pennebaker, Groom, Loew, & Dabbs, 2003). Overall, testosterone had the effect of suppressing the participants' use of non-I pronouns. As testosterone levels dropped in the weeks after the hormone injections, the participants began making more references to other people. No consistent mood or other language correlates of testosterone emerged. The study suggested that one function of testosterone, then, may be to steer people's interests away from other people as social beings.

Word count tools have also provided novel ways of assessing status. In natural conversation in the lab, in emails, and even in the Watergate tapes between Nixon and his aids, the relative use of 1st person singular pronouns has been found to be a robust marker of the status of two people in an interaction (Chung & Pennebaker, 2007; Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2009; Niederhoffer & Pennebaker, 2002). Specifically, people tend to use more 1st person singular pronouns when lower in status relative to their audience. This greater degree of self-focus is likely needed for lower status people to monitor their behavior to accommodate their higher status interaction partners. Indeed, across days of email exchanges, it has been found that people consistently regulate their levels of self-focus, as measured through "I" use, depending on the relative status of their correspondent (Kacewicz et al., 2008).

As demonstrated above, word use can reflect the changing goals of speakers in relation to their audience. Another example is in the examination of the press conferences of New York City's old Mayor, Rudolph Giuliani (Pennebaker & Lay, 2002). During the first four years of his administration, Giuliani used a very high rate of "we" words and negative emotion words, and a low level of "I" words and positive emotion words. Soon

after his diagnosis of prostate cancer and the breakup of his marriage, his personality was widely reported as changing. He became warmer, more sensitive to others, and more genuine. Giuliani's apparent personality shift was associated with large increases in "I", drops in "we", and modest increases in positive emotion words. After 9/11, his language shifted again, with an increase in "we", and increases in both positive emotion and negative emotion words. These linguistic changes were accompanied by changes in the ways that people in his office and the general public perceived him to be.

One example of word use reflecting group processes is in Sexton and Helmreich's (2000) analysis of the cockpit communication of dozens of flight teams that took part in a NASA B727 simulator study. They found that pronouns varied as a function of how long the crew had worked together; the longer the group had worked together, the less people used I-words and the more they used we-words. The dropping-I, increasing-we phenomenon is a reliable effect apparent in the transcripts from a large therapy group (N=22) that met for a week every 6 months for three years (Odom, 2006), in the lyrics written by the Beatles over their 10-year collaboration (Petrie, Pennebaker, & Sivertson, 2009), and in an analysis of the Watergate tapes (Mullen, Chapman, & Peaugh, 2001). Interestingly, in the latter study, I-use was found to decrease as a function of the number of discussants (3, 4, or 5), which the authors suggested was indicative of a decrease in self-focus within larger working groups.

As found in the above studies, "we" use can indicate relationship closeness. Indeed, previous studies have shown that using "we" in interactive tasks in the lab predict relationship functioning and marital satisfaction (Gottman and Levenson, 2000; Sillars,

Shellen, McIntosh, & Pomegranate, 1997; Simmons, Gordon, and Chambless, 2005). However, a study of over 80 couples interacting outside of the lab with each other via IM over 10 days failed to show a relationship with 1st person plural pronouns. Rather, the more that participants used emotion words in talking with each other – both positive and negative emotion words – the more likely their relationship was to survive over a 3 to 6 month interval (Slatcher & Pennebaker, 2006).

The research suggests that although brief speech samples can be reliably related to the functioning and quality of a relationship, natural language outside of the lab can provide a different picture of what types of communication patterns are associated with long-term relationship stability (see also Fellows & Pennebaker, 2008). Considering these findings, and considering that IM is now the preferred method of on-line communication (Shiu & Lenhart, 2004), there is a clear need for future research to consider word use in naturalistic computer-mediated communication.

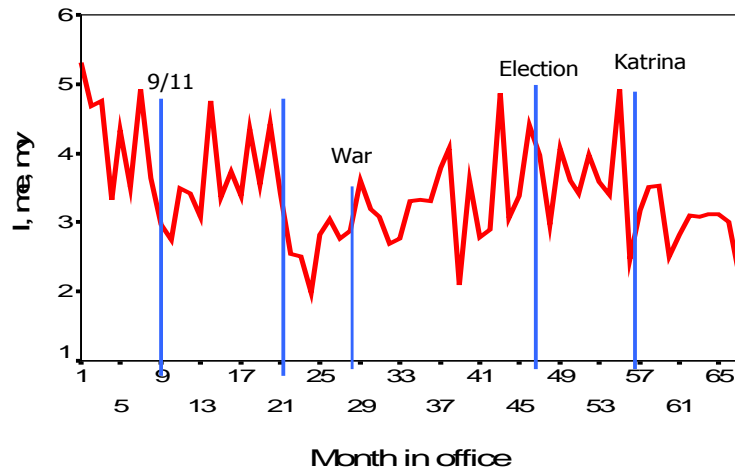
Tracking Cognitions over Time

In tracking the public statements by two of al-Qaeda's top leaders Bin Ladin and al-Zawahiri over time, it was possible to get a glimpse into their personalities and psychological states that were otherwise impossible to acquire (Pennebaker & Chung, 2008). Initially, the 58 translated al-Qaeda texts representing statements from the 1980s to 2005 were compared with those of other terrorist groups from a corpus created by Smith (2004). The al-Qaeda texts contained far more hostility as evidenced by their greater use of anger words and 3rd person plural pronouns.

As for the individual leaders' language use over time, bin Ladin evidenced an increase in the rate of positive emotion words as well as negative emotion words – especially anger words. He also showed higher rates of exclusive words (e.g. but, except, exclude, without) over the last decade, which often marks cognitive complexity in thinking. On the other hand, al-Zawahiri tended to be slightly more positive and significantly less negative and less cognitively complex than bin Ladin in his statements. He evidenced a surprising shift in his use of 1st person singular pronouns over the last two years. This was interpreted as indicating greater insecurity, feelings of threat, and, perhaps, a shift in his relationship with bin Ladin.

Another example is in President George W. Bush's use of 1st person singular during his (over 500) formal and informal press meetings over the course of his presidency (Pennebaker, 2009, unpublished data, see Figure 1). Note that only the press interactions where Bush was speaking off the cuff rather than reading any prepared remarks were analyzed. As can be seen in the graph below, there were large drops in I-usage immediately after 9/11, and again with Hurricane Katrina. Most striking were the drops starting in August 2002 – just prior to the Senate's authorizing the use of force in Iraq. Perhaps this was when Bush formally decided to go to war in Iraq, an event that caused attention to be focused away from the self. If true, one can begin to appreciate how word counts can betray intention and future action.

Figure 1: President George W. Bush's Use of First Person Singular Pronouns during his Months in Office.



The analysis of the natural language of these political leaders highlights the ability of computerized word counts to reveal how people are attending and responding to their personal upheavals, relationship changes, and world events. An analysis over time provided a within subject analysis that bypassed the difficulties in traditional self-reports (i.e. personally smoking out these leaders in their top secret hideouts to ask them to fill out questionnaires with minimal response biases).

Another major area in which changes in cognitions are tracked over time is in psychotherapy sessions, where the goal is to change maladaptive thoughts and behaviors. Mergenthaler (1996) used word counts to identify word categories that characterize key moments in therapy sessions in order to provide an adequate theory of change. He found that key moments of progress are characterized by the co-occurrence of emotion terms and abstractions (i.e. abstract nouns that characterize the intention to reason further about that term) in a case study and in a sample of improved vs non-improved patients. These

suggest that being able to express emotions in a distanced and abstract way is important for therapeutic improvements.

The text analysis programs used by these clinicians, such as Bucci's Discourse Attribute Analysis Program (DAAP; Bucci & Maskit, 2005), are similar to LIWC in that they use a word count approach, and many of their dictionary categories are both grammatical and empirically derived. However, the grammatical categories for the clinical dictionaries are broad (e.g., they throw all function words into a single category), and their empirically derived categories are based on psychoanalytic theories and clinical observations.

More recently, Wolf, Chung, and Kordy (in press) inductively derived clusters of co-occurring words in 4,241 e-mails between 297 individuals and their psychotherapists during an aftercare program following inpatient psychotherapy. Each of the nine clusters represented coherent themes centered on physical symptoms, coping, adjustments at work, family, relationships, treatment, exercise and diet, leisure activities, and household issues. The authors found that inpatients who had improved the least during inpatient psychotherapy tended to use more words in the physical symptoms cluster in their emails than did patients who evidenced therapeutic gains during inpatient treatment. These results suggest that unsuccessful inpatients carry the continued burden of managing symptoms that interfere or prevent them from starting to tackle adjustments to other life domains. Their study showed how inductively derived word clusters could be adapted to online aftercare settings for which psychoanalytically derived word count categories may not be the most suitable approach.

The advantage of all word count tools for the analysis of therapeutic text is that word counts tend to be a less biased measure of therapeutic improvements than clinician's self-reports (Bucci & Maskit, 2007). In addition, word count tools can be assessed at the turn level, by sessions over time, and for the overall treatment, making word count approaches a powerful tool for assessing micro processes of self-regulation, and larger processes in self-change.

SUMMARY

As reviewed in the above studies, the study of Internet natural language text has allowed the psychological study of word use to extend beyond lab studies in many ways. Language use has been reliably linked to individual differences, psychological states, and psychological processes over time. By analyzing natural language using computerized text analysis tools, researchers have found that use of function words are markers of psychological states across a wide variety of texts, and that reliable changes in function word use can signal changes in psychological states. This would have been extremely difficult to discover using manual coding, since function words are often "invisible" to readers of natural language text. The research reviewed above suggests that word counts of content word categories over time can also indicate changes in psychological states. For example, the research on word counts in therapeutic contexts highlights the ability of word count tools to provide a probabilistic but fair diagnosis of therapeutic improvements (Bucci & Maskit, 2007; Wolf et al., 2009).

The next step in word count research is to establish the predictive validity of word counts for psychological processes. One way to showcase the ability of word count

methods to capture psychological processes over time and to predict important outcomes is to examine personal blogs devoted to self-change. The research reviewed above suggests that examining both function and content words using computerized text analysis can provide an assessment of psychological states throughout the self-change process.

Chapter 2: The Present Study: Predicting Weight Loss Based on Natural Language in Blogs

While previous studies have examined questionnaires completed by participants, either retroactively or as structured daily reports, no known studies have examined open-ended public diaries about the weight loss process. The current study presents an improvement over previous methods of examining the process of self-change. Blogs document the subjective experience of self-change almost daily for long periods of time. There are observable and quantifiable features of blogs, such as blogging rate and comments that can be used to measure psychological constructs relevant to the self-change process, such as self-monitoring and social support. In addition, language use can be measured using LIWC, allowing a comparison to previous and future studies using LIWC.

For the current study, all blogs from one of the largest dieting communities on the web, www.dietdiaries.com (2007) were harvested. Most therapies assume that simply putting thoughts and feelings into words through talking or writing can bring about change (Pennebaker, 1997), and so examining language use in blogs might help to uncover the possible mechanisms of self-change. Using LIWC, evidence for each of the primary self-regulatory strategies found to be linked to successful weight loss in lab and clinical based studies will be linked to body weight change over the course of blogging. Theoretically, the processes underlying successful weight loss in bloggers should be similar—or at least involve the same general types of processes—to how weight loss occurs in lab and clinical based studies (Weinstein, 2006). According to previous

research, the words to be examined should tap the various strategies found in the micronarratives, with support from lab and clinical studies. Possible correlates of weight change are:

SELF-FOCUS: FIRST PERSON SINGULAR PRONOUNS

The high use of first person singular pronouns has been linked to self-focus. Indeed, the use of 1st person singular has been found to be a better marker of depression, characterized by a high degree of self-focus and rumination, than use of negative emotion words (Mehl, 2006b; Rude, Gortner, & Pennebaker, 2004; Stirman & Pennebaker, 2001; Weintraub, 1989). High degrees of self-focus have been found to be aversive and lead to negative affect even in healthy college student populations (Duval & Wicklund, 1972). The self-regulation literature suggests that when dieters have high degrees of self-focus, negative self-schemas are activated regarding their weight, body image, and overall self-worth (Heatherton & Baumeister, 1991). This negative self-focus has been shown to lead to overeating in dieters across multiple studies (e.g. Beebe, Holmbeck, Albright, Noga, & DeCastro, 1995; Heatherton, Striepe, & Wittenberg, 1998; Schupak-Neuberg & Nemeroff, 1993; Wallis & Heatherington, 2004). It has been theorized that binge eating can be viewed as an escape from self-awareness, when high degrees of self-focus lead one to narrow attentional resources to the immediate environment (Heatherton & Baumeister, 1991).

Findings regarding self-focus and pathological eating styles have been found in Internet language use by pro-anorexics (i.e. anorexics who espouse a lifestyle of starvation). Using self-references as a measure of self-focus, Lyons, Mehl, and

Pennebaker (2006) found that pro-anorexia online message boards contained fewer self-references, fewer cognitive mechanism words, and more positive emotion words than online message boards for recovering anorexics. The authors suggested that attention away from the self allowed pro-anorexics to stabilize emotionally, and to sustain their immunity to psychological treatment.

While a high degree of self-focus has been linked to the activation of negative self-schemas and to overeating in dieters, the research on pro-anorexics suggests that a moderate degree of self-focus may be required for successful self-change. Indeed, most psychotherapies are based on the idea that self-reflection is important in the self-change process. Given these competing views, no specific hypotheses were made about the relationship of first person singular pronouns (e.g. *I, me, mine, my, myself*) and weight loss. First person singular pronoun use was measured using LIWC2007.

EMOTION WORD USE

Emotions, and especially negative emotions, have been linked to disinhibited and binge eating through the use of retrospective self-reports of mood (Eldredge, Agras, & Arnow, 1994), and mood inductions in the laboratory (Cools, Schotte, & McNally, 1992; Ruderman, 1985). While eating in response to emotions has been hypothesized to be a coping strategy, recent research suggests an alternate theory. The alternate theory states that self-regulation is a limited resource of attention (Baumeister & Heatherton, 1996; Heatherton & Baumeister, 1996). When attentional resources are depleted upon inhibiting negative emotions, a dieter has fewer attentional resources to devote to inhibiting food intake, possibly leading to overeating or breaking a diet.

In support of the limited resource theory, Vohs and Heatherton (2000) found that chronic dieters ate significantly more ice cream in response to a sad movie if told to inhibit their emotions, relative to chronic dieters who were told to express their emotions naturally. Both groups experienced increased distress in response to the movie, but only the group that inhibited their emotions was less able to inhibit their ice cream intake. The findings could not be explained in terms of a coping model. Instead, the findings supported their theory that self-regulation is a limited resource that enables one to successfully control impulses and desires (Baumeister & Heatherton, 1996; Heatherton & Baumeister, 1996). Specifically, when self-regulatory resources are “used up” on controlling or inhibiting one activity, these resources are depleted, and unavailable for controlling or inhibiting another activity. According to this theory, then, the expression of emotions is an important strategy in regulating behavior in other domains, such as dieting.

In the current study, the degree of negative emotion word use (e.g., *angry*, *anxiety*, *cry*, *kill*, *sad*, *ugh*) and positive emotion word use (e.g., *excitement*, *hug*, *lol*, *love*, *yay*) was measured by LIWC2007. It was hypothesized that greater emotional expression would lead to freeing of attentional resources for successful dieting.

COGNITIVE WORD USE

The micronarrative research using the National Recovery Dictionary suggested that cognitive evaluations, as indicated by words such as *belief*, *commit*, *consider*, and *realize* are important determinants in successful self-change from substance abuse (Sobell et al., 2001). The increasing use of cognitive mechanism words (e.g. *caused*, *insight*,

think, understand) in written narratives about traumatic or stressful events has also been found to be an important predictor of improved physical health, fewer physician visits, and better immune functioning, among a host of other salutary effects (Low, Stanton, & Danoff-Burg, 2006; Pennebaker, Mayne, & Francis, 1997). The use of cognitive mechanism words suggests that the writer is demonstrating flexibility in thinking about a past negative event in order to make sense of a negative event, freeing cognitive resources for other activities (Boals & Klein, 2005; Klein & Boals, 2001a; 2001b).

The use of cognitive mechanism words, then, might be seen as increasing the capacity of attentional resources for other activities, such as dieting. Therefore, it was hypothesized that a greater use of cognitive mechanism words in blogs would be associated with more weight loss. The use of cognitive mechanism words were measured using LIWC2007's cognitive mechanisms category.

SELF-MONITORING OF EATING BEHAVIOR

Ingestion Word Use

Wing and Hill (2001) found that keeping records of dietary intake was an important determinant of successful weight loss in weight loss programs. Similarly, in a review of clinical weight interventions, Weinstein (2006) found that weight loss success was positively correlated with self-monitoring (i.e. consistent records of dietary intake) and with attendance at weight loss meetings. In a more recent study of a diverse group of overweight and obese adults in a behavioral weight loss program, more frequent attendance at group sessions, more food diaries per week, and more exercise were

predictive of greater weight loss (Hollis, Guillon, Stevens, Brantley, Appel, Ard et al., 2008).

One large prospective study further found that self-monitoring was key to success in weight loss maintenance (Gorin, Phelan, Wing, & Hill, 2004). The National Weight Control Registry (NWCR) recruits participants who have gained no more than 5 lbs for at least 1 year after losing at least 30 pounds. The study did not examine weight losers. In the year that followed entry in the NWCR, successful weight loss maintainers, relative to weight gainers, were generally more consistent in their dieting behaviors over time. They continued to ingest fewer calories overall and fewer calories from fat, continued engaging in physical activities, weighed themselves frequently, ate regular meals, including breakfast, and reported exerting the same degree of dietary restraint on weekdays as weekends and on holidays and vacations as the rest of the year. Keeping a careful and consistent eye on food intake then, seems key to successful weight loss maintenance.

In the current study, food monitoring was measured by the degree to which the ingestion word categories (e.g., *binge, carbs, eat, pasta, vodka*) appear in the blogs. Ingestion words were measured by LIWC2007. Since lab and clinical studies have found the keeping of food intake diaries to be associated with weight loss, it was hypothesized that a high use of ingestion words would be associated with greater weight loss.

Rates of Blogging

The very act of reporting on the “little experiences” of everyday life was first termed “self-monitoring” by behaviorists for behavioral modification (see Wheeler & Reis, 1991). Behaviorists, who were concerned primarily with objective counts of

behavior, suggested that the mere recording of a behavior could change its frequency precisely because of its reactivity. As such, daily behavioral diaries are used by behaviorists to report on target behaviors and surrounding events (see Lewinsohn & Talkington, 1979).

Since the frequency and rates of blogging about one's weight loss are records that one has monitored or reported on one's own behavior, it was hypothesized that a greater frequency and rate of blogging would be associated with greater weight loss. In the current study, the rates of posts by a blogger were computed by total word count, the number of entries posted per week, and the number of active weeks (i.e. the number of weeks for which there was at least one entry).

SOCIAL SUPPORT

Social support has been found to be an important determinant in weight loss success. A meta-analysis of 13 studies that looked at the effect of spousal involvement in weight loss found a moderate effect in the initial months of treatment (Black, Gleser, & Kooyers, 1990). One study found that participants recruited into a weight loss treatment program with friends lost more weight than did participants who were recruited alone (Wing & Jeffery, 1999). The degree of weight loss was most similar between participants and friends in the social support condition. Similarly, a later study found that the weight loss success of partners (not the number of partners) was a strong predictor of participant weight loss at 6, 12, and 18 months into treatment. Overall, partners positively influenced the number of food diaries completed, and adherence to the treatment program (Gorin, Phelan, Tate, Sherwood, Jeffery, & Wing, 2005).

Taylor (2007) operationalized social support and social integration as “the number of relationships or social roles a person has, the frequency of contact with various network members, and the density and interconnectedness of relationships among the network members.” (p.146). Social support within blogs can be observed through the comments or replies to a blog entry, and through the number of different people who respond to a blog or comment on a blog. Feedback from a blog audience through comments makes up approximately 30% of the blogosphere (Vigas, 2005). In examining a blog corpus, Mishne and Glance (2005) found that for those days in which a blog post received a comment, there was a mean of 6.3 comments ($SD = 20.5$). In a sample of bloggers, Vigas (2005) found that 20% of bloggers reported that the comments people post on their blog sites affect what the bloggers write about, but that bloggers still reported disclosing highly personal topics. Furthermore, research has shown that the kind of feedback received is important: McKenna and Bargh (2002) found higher rates of participation by members of an online community when they received positive feedback than when they received negative feedback by other group members.

In this study, social support was measured in various ways. The number of comments left by other bloggers to each blog entry and the number of different commenters were counted. These were considered to be measures of the responsiveness of the online social network to each blog. In addition, the degree of positivity and negativity in the responses for each blogger were measured using LIWC2007’s positive emotion words and negative emotion words categories. Within the blogger’s posts themselves, the degree of social word use (e.g., *family*, *greetings*, *husband*, *sharing*, and

pronouns such as *she*, *they*, and *you*, indicating an interest or awareness of others) were assessed using LIWC2007 to measure the degree to which the blogger was attending to their social world. In addition, the number of comments made by a blogger to other blogs in the community, and the number of different blogs commented to were counted as a measure of investment in the blog community. It is hypothesized that a high use of social words in the blogs, a greater number of responses, more positivity in responses, and less negativity in responses, will be associated with greater weight loss. Also, high rates of both receiving and making comments were expected to be associated with greater weight loss.

SUMMARY

The goal of this dissertation was to examine the strategies people use in blogging about weight loss, and to establish a prediction model of successful weight loss based on computerized word counts of the blogs. The goal of the study was not to say that language causes weight loss or weight gain, but that language is reflective of what people are attending to, and how they are thinking (Slobin, 1996) during attempts at self-change. Examining language use in blogs through computerized text analysis presented a non-obtrusive and efficient way to examine the underlying processes in attempts at self-change. In addition to examining the language used in blogs, the rate of blog entries, along with the comments received and made to the community were examined to assess the degree to which self-monitoring and social support influenced weight loss within an online community of diet blogs. Also, the degree to which starting weight and the magnitude of the goal (or goal discrepancy from starting weight) influenced the degree of

weight loss achieved were assessed. Presumably, more realistic goals would be easier to achieve than unrealistic goals (Herman & Polivy, 2003). On the other hand, having larger goals might reflect one's dedication to blogging and/or dieting, and so there was a possibility that larger goals would be associated with more weight loss.

An entire community of public blogs devoted to weight loss (DietDiaries; www.dietdiaries.com, 2007) was examined. A sample of blogs that had entries spanning at least 15 weeks was selected. A period of 15 weeks was chosen since this was the mean duration of weight loss studies in a 1997 meta-analysis of all studies on therapeutic weight loss treatments for obesity published in peer-reviewed scientific journals over a 25 year period (> 700 studies; Miller, Koceja, & Hamilton, 1997). Weight loss during a 15-week period in the meta-analysis produced a weight loss of about 11 kg (i.e. 24.2 lbs), and so it was possible to compare the degree of weight loss in therapeutic treatments to weight loss in blogs. Changes in language use were tracked along various dimensions: emotions, cognitive processing, self-references, social words, and eating and food related words. Responses to the blogs and the rates of posting entries were examined as markers of social support and self-monitoring, respectively. In addition, starting and goal weight discrepancy from starting weight were examined.

Three primary research questions were assessed using a word count approach, along with counts of entries and comments.

Research Question 1: What Predicts Weight Loss in Blogs?

Examining rates of word use in blogs can reveal how much a person is attending to the world along various dimensions. For example, using high rates of positive emotion

words suggest that the person is thinking along a positive dimension, using high rates of first person singular pronouns suggests that the person is attending a lot to themselves, using high rates of social words suggests that the person is attending to their social worlds, etc. Therefore, measuring the overall uses of these categories throughout the 15 week period and correlating them with overall weight change over the 15 week period might reflect how attending to various dimensions is associated with successful self-change.

Social support was assessed based on the use of social words, by the number of comments by other bloggers to the blog entries, and by the number of comments made to other bloggers in the community. Social support was also assessed by counting the number of different people that commented to a blog, and the number of different blogs within the community that a blogger commented to. These were considered as measures of social support and investment in the diet blog community. The rates of blogging were also quantified through word counts, counts of entries per week, and counts of active weeks (i.e. number of weeks for which there was at least 1 entry on a blog). The magnitude of the goal was assessed by subtracting goal weight from starting weight.

It was hypothesized that greater social support and rates of blogging would be predictive of weight loss success. A high rate of emotion words, cognitive mechanism words, and social words were expected to be associated with weight loss, and no predictions were made for the use of first person singular pronouns, or for starting and goal weight information. The linguistic dimensions, rates of blogging, weight

information, and social support measures were entered into a regression equation to predict percentage of body weight change over the 15 week period.

Research Question 2: Do Changes in Word Use over Time Lead to Weight Loss?

Just as there are changes in expressive writing over three days of writing that are predictive of improved health (Pennebaker et al., 1997), perhaps blogging about weight loss leads to changes that are predictive of actual weight loss. The first half and the second half of blogs during the first 15 weeks of blogging were compared. Specifically, difference scores in the language use, rates of blogging, and measures of social support were computed. These difference scores were correlated with percent body weight change. A regression analysis using the difference scores was conducted to predict percent body weight change.

It was predicted that an increasing use of cognitive mechanism words would be predictive of greater weight loss, and that a large drop in blogging rates and social support measures would be predictive of weight gain. Since self-monitoring has been found to be an important determinant of weight loss and weight loss maintenance, it was predicted that drops in rates of ingestion word use, in word counts, in mean entries per week, and in the number of active weeks would be associated with less weight loss or weight gain. Similarly, since social support has been found to be associated with weight loss in weight loss programs, it was predicted that a drop in social support (i.e. in comments received or made, in the number of different bloggers who commented or were commented to by a blogger, in social word use, and in positive emotion word use in the comments), would be associated with less weight loss.

Research Question 3: Can We Predict Dropping Out or Weight Loss from the First Two Entries?

Norcross, Ratzin, and Payne (1989) followed people who had made a New Year's Resolution related to behavioral change (e.g. weight loss, smoking cessation, exercise, etc.). They found that only 77% of resolvers kept their resolutions continuously for one week, 55% for 1 month, and only 19% at a two-year follow-up (Norcross & Vangarelli, 1989). Readiness to change and self-efficacy were strong predictors of sticking to the resolution at 1 week and 1 month follow-up. Note that 33% of their sample did not keep their resolutions for beyond one week, and about half of their sample had quit their resolutions for behavioral change at 1 month.

In order to identify the start strategies of dropouts (i.e. those who quit blogging after only two posts), and dedicated bloggers (i.e. those who blog for at least 15 weeks), the first two entries of each blog were examined. For the first two entries, measures of language use, social support, and the latency to blog the 2nd entry were correlated with percent body weight change observed for the entire duration of blogging. A binary logistic regression analysis was conducted to determine if these measures based on only the first two entries could predict group membership (i.e. dropouts vs. dedicated bloggers).

The literature suggests that greater social support would be associated with continued or more frequent participation on the blog, and so it was predicted that social support would be an important predictor of dedicated blogging. Specifically, it was predicted that making many comments to and receiving many comments by different

bloggers would be indicative of the level of investment in the diet community, and that these would be associated with a greater likelihood of being a dedicated blogger. In addition, since frequency of participation on an online community has been associated with receiving positive feedback, a high rate of positive emotion words in the comments received during the first two entries was predicted to be associated with a greater likelihood of being a dedicated blogger, while a high rate of negative emotion words in the comments was predicted to be associated with a decreased likelihood of being a dedicated blogger.

In order to assess whether weight loss could be predicted by the first two entries, a linear regression was conducted on the target variables based on the first two entries of each blogger. It was expected that the predictors of weight loss for Research Question 1 would be similar to those found for the first two entries.

Chapter 3: The Weight Loss Blogs Corpus and Sample

The goal of the dissertation was to predict weight change in blogs devoted to weight loss using a word count approach. A sample of blogs was selected from a corpus of blogs devoted to weight loss according to various criteria. A blog is the collected work (i.e. all written “entries” or “posts” across time) of a single person or “blogger”. LIWC was used to assess word categories in the blogs underlying various known self-regulation strategies, such as self-focus, emotional expression, cognitive processing, self-monitoring through food diaries, and social support. These constructs were examined over the course of several months of blogging using a word count tool. In addition, the rates of blog posts and comments to the posts were assessed as measures of self-monitoring and social support, respectively. The current chapter examines the characteristics of the weight loss blogs corpus.

WEIGHT LOSS BLOGS CORPUS

A public community of blogs devoted to weight loss, www.dietdiaries.com, began in 1999. The corpus of all publicly available blogs (2,637 blogs; 46,242 entries available to anybody with Internet access without a login or password for the site) since 1999 was harvested in November 2007. Any blogs indicating that the blogger was less than 18 years of age at the start of the blog were deleted ($n = 107$), and therefore not included in any analyses in order to examine dieting in adults only. This resulted in a total of 2530 blogs. Anybody interested in blogging about their weight loss goals was free to join the community by starting a blog at any time, meaning bloggers within the www.dietdiaries.com community had different start dates and different durations of

blogging. Although bloggers were not required to post how they had come across the www.dietdiaries.com website, many bloggers spontaneously reported that they had come across the community by browsing online forums devoted to weight loss, such as Yahoo! Groups. Below is their description of the community from the homepage (DietDiaries; www.dietdiaries.com, 2007):

We are a supportive community of individuals with like minded goals of becoming healthier. You can choose to start your own diet diary today or just read other's entries for inspiration. For those seeking support, we suggest starting a "public" diary. A "public" diary can be read by others and also allows other members to leave comments of encouragement as they pertain to your entries....

Information from each public blog entry was automatically downloaded into a database with the date of the entry, and the blogger's self-reported weight for each entry, which was available for most entries. Each blog was assigned a five digit number in order to uniquely identify the blog. Most bloggers updated their weight with each entry. There were 2 false records (i.e. 2 blogs of blank pages), and 202 blogs without any words written within any of the entries (i.e. these blogs had been set up with entries with date stamps, and perhaps some information in the blog profile or weight recorded in the entries, but there were no written entries for these blogs). Below is a table with descriptive information on the entire corpus of www.dietdiaries.com blogs.

Table 1: Descriptive Information on www.dietdiaries.com Corpus

Measure	<i>N</i>	<i>M</i>	<i>SD</i>	Median	Min	Max
Age in Years	1331	31.18	10.23	29.00	18	78
Blog Entry Rates						
# Entries per Blog	2530	17.96	51.55	3.00	1	781
# Days per Blog	2530	107.30	296.44	6.00	1	3116
# Weeks per Blog	2530	15.93	42.30	1.00	1	446

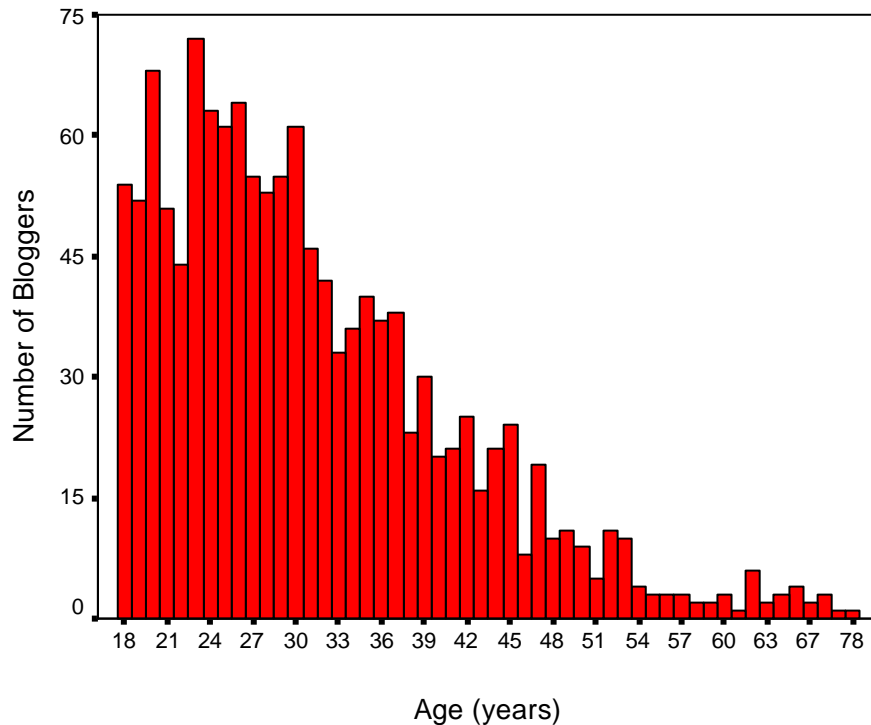
Total Word Count per Blog	2530	4,959.49	20,300.89	493	0	387,889
Word Count per Entry	2530	185.11	196.60	132.83	0	2,708
Weight Loss Goals						
Start Weight in lbs	2045	202.47	54.42	193.00	84	450
End Weight in lbs	2014	197.92	53.14	189.50	84	590
Goal Weight in lbs	1675	141.63	23.96	141.63	75	270
Goal # lbs to Lose	1482	62.92	43.89	50.00	1	253
Change in Weight in lbs (end weight – start weight)	1866	-4.67	13.92	0.00	-148	71
Percent change in body loss	1866	2.06	5.76	0.00	47.78	33.33
Social Support						
Total comments received	2530	45.04	170.35	5.00	0	3576
Comments received per entry	2530	1.89	1.60	1.57	0	13
Total comments made to other blogs	1130	71.63	248.51	7.00	1	2979

Sex and Age

Demographic information was available from either the brief description page, or from reading the posts (using cues such as birthday ages, explicit mentions of age, pronoun use by commenters, descriptions of body parts, menstrual cycles, or relationship roles – e.g. boyfriend, daughter, mother, husband, etc.). It was possible to determine the sex of 1848 of the 2530 bloggers. Of these 1848 bloggers, 1773 (i.e. 95.94% for those in which sex was known) were female.

Age information was available for 1331 of the 2530 blogs. Of these 1331 bloggers, the average age was 31.18 ($SD = 10.23$) years old (see Figure 2).

Figure 2: Histogram of Age at Start of Blog for Bloggers in the Corpus where Starting Age Information was Available (n = 1331).



Blogging Rates

Number of entries

On average, bloggers had a mean of 17.96 entries ($SD = 51.55$). The median number of entries was 3.00, and the modal number of entries was 1.00. Approximately 1/3 of the bloggers (30.12%; 762 bloggers) posted only one entry. Half of the bloggers (50.99%; 1290 bloggers) posted 3 entries or less. Approximately eighty percent (79.68%) of the bloggers ($n = 2016$) posted 15 entries or less (see Table 2).

Table 2: Number of Blogs with 1 to 15 Entries

Total Number of Entries	Number of Blogs	Cumulative Proportion of Corpus with Valid Entries ($n = 2530$)
1	762	30.12
2	329	43.12
3	199	50.99

4	146	56.72
5	140	62.25
6	81	65.45
7	67	68.10
8	43	69.80
9	39	71.34
10	39	72.89
11	48	75.78
12	40	76.36
13	31	77.59
14	31	78.81
15	22	79.68
Total	2016	79.68

As can be seen in Table 3, there were a dozen bloggers who blogged over 400 entries.

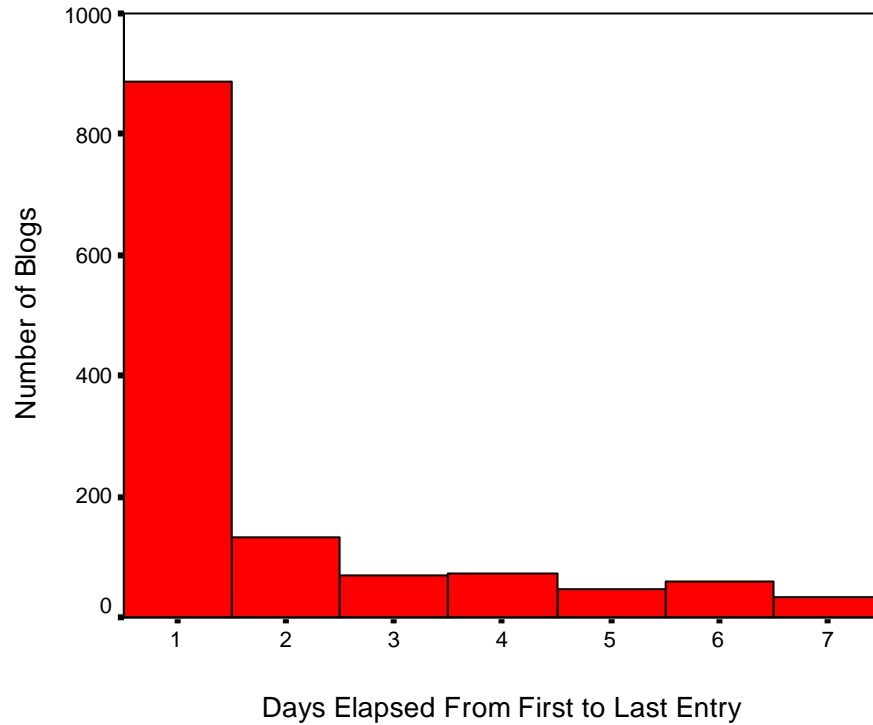
Table 3: Number of Blogs with 16 Entries or More

Total Number of Entries	Number of Blogs	Cumulative Proportion of Corpus with Valid Entries (n = 2530)
Up to 100	2433	96.17
101 to 200	55	98.34
201 to 300	19	99.09
301 to 400	11	99.52
401 to 500	7	99.80
501 to 600	4	99.96
601 to 700	0	99.96
701 to 800	1	100.00
Total	2530	100.00

Duration of Blogs

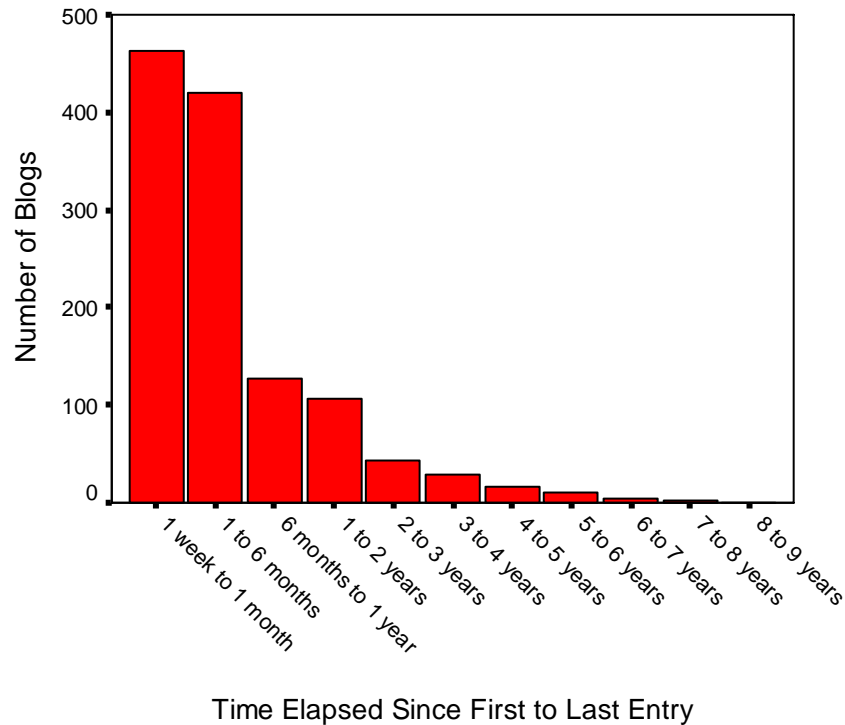
About 1/3 of the bloggers (n = 888; 35.10%) blogged for only one day. Half of the bloggers (51.58%; 1305 bloggers) kept their diet blog for a week or less (see Figure 3). The mean number of days in which the blogs contained an active entry since the beginning of the blog site to the data harvest date was 107.30 days ($SD = 296.44$). However, this range varied from 1 day to 3116 days (this is equivalent to 445.14 weeks; approximately 102.38 months; or 8.53 years).

Figure 3: Histogram of Total Number of Posts by Bloggers in the Corpus Who Blogged One Week or Less (51.58% of all Bloggers)



Considering that a third of the bloggers posted only once, the median duration of diet blogs was brief; 6.00 days. Surprisingly, there was 1 blogger who had kept their blog for over 8 years, 2 bloggers who had kept their blog for 7 to 8 years, and 5 bloggers who had kept their blog for 6 to 7 years (see Figure 4).

Figure 4: Histogram for Blogs with Durations Greater than One Week



Word Count

The mean number of words per entry was 200.92 (*SD* = 196.92) words. The mean total word count for a blog was 4,959.49 (*SD* = 20,300.89). Although there were quite a few entries (*n* = 202) with no words written (i.e. these were entries with only weight information, date, etc., but with no typed words), there was one blog entry with a total of 3,104 words.

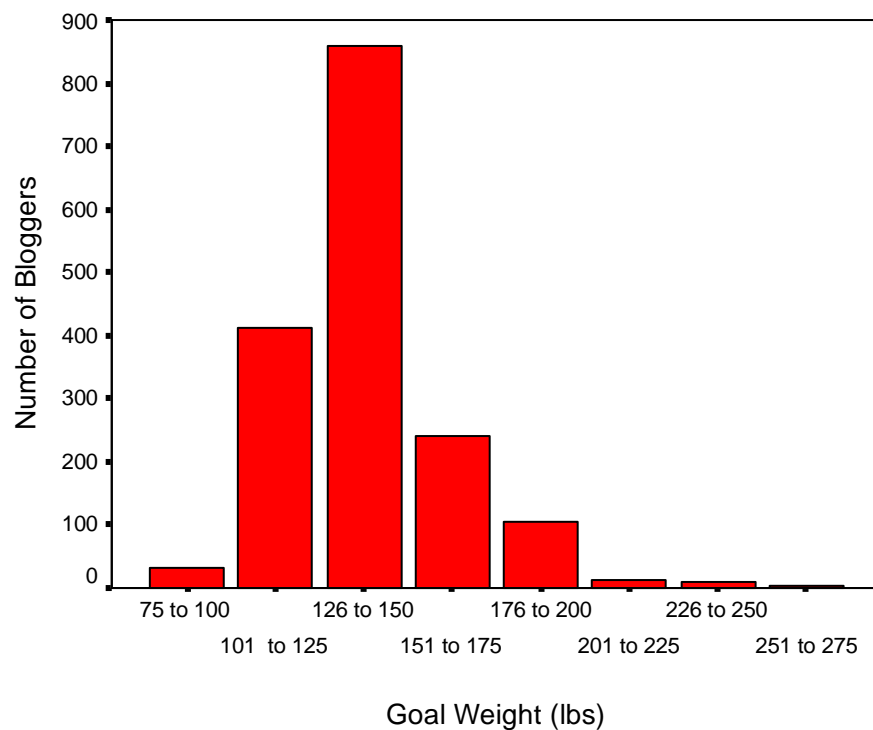
Weight Information

Goal Weight

In order to determine goal weight for each blogger, goal weight information was manually gathered for each blogger by examining the brief description page, by reading the first few posts of the blog, or by examining a commonly used weight ticker (a picture

of a scale indicating current weight, start weight, and goal weight). Any weight information in the blogs was converted to pounds. Thirty-five bloggers (1.38%) had reported their weight in kilograms (1 Kg = 2.20 lbs). Goal weight at the start of the blogs was determined for 1675 bloggers (see Figure 5). The average goal weight of these bloggers was 141.63 lbs ($SD = 23.96$).

Figure 5: Histogram of Goal Weights in Pounds for those Blogs in which Goal Weight Information was Available (n = 1675)



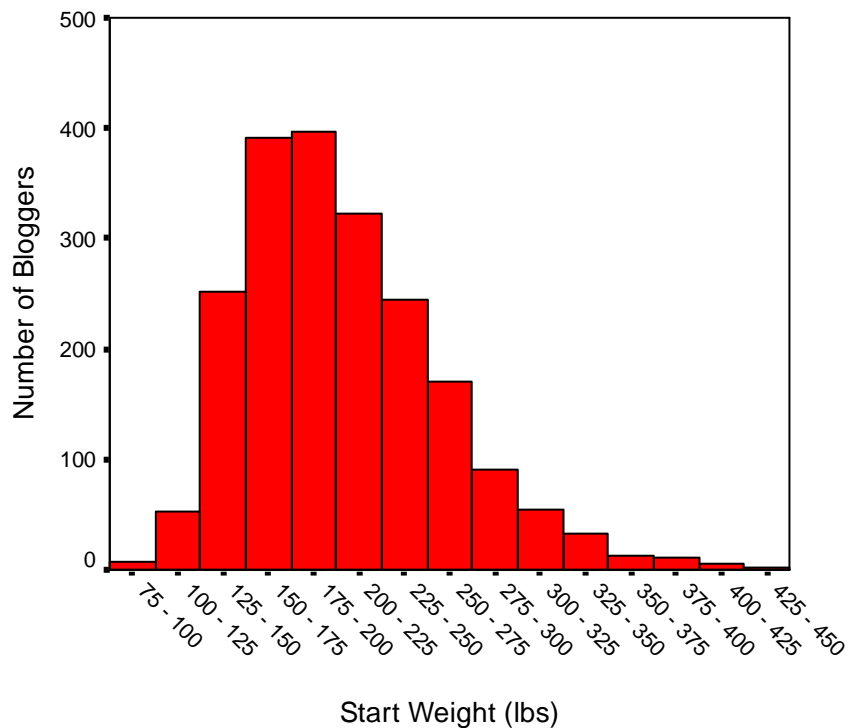
Starting Weight

Starting weight was determined by the first recorded weight in the blog. For most bloggers (n = 1325), this information was recorded in the first entry, for some bloggers, this information was taken from their second or third entry (n = 720). Otherwise, starting

weight was listed as unknown ($n = 485$). The average starting weight of the bloggers was 202.47 lbs ($SD = 54.42$; see Figure 6). Blogs ($n = 446$) indicating a start weight less than the lowest desired goal weight (75 lbs) were excluded since these suggested a typing error or no starting weight entry.

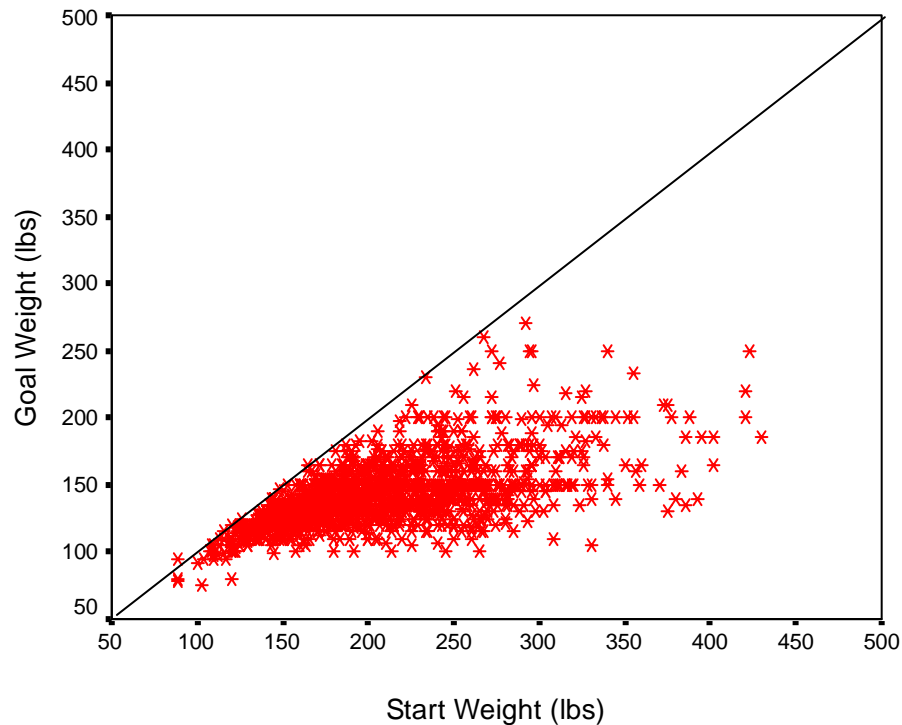
Recall that the average goal weight was 141.63 lbs, representing an average desired decrease of 62.92 lbs ($SD = 43.89$). On average, bloggers lost approximately 2% ($SD = 5.76$) of their body weight over the course of blogging.

Figure 6: Histogram of Starting Weights for Bloggers (i.e. Weight at First Entry; $n = 1325$)



As can be seen in Figure 7, the goal weights tended to not be too far off from the starting weights. However, it appeared that bloggers with starting weights over 300 lbs had a broader range of goals, and much more difficult goals to achieve.

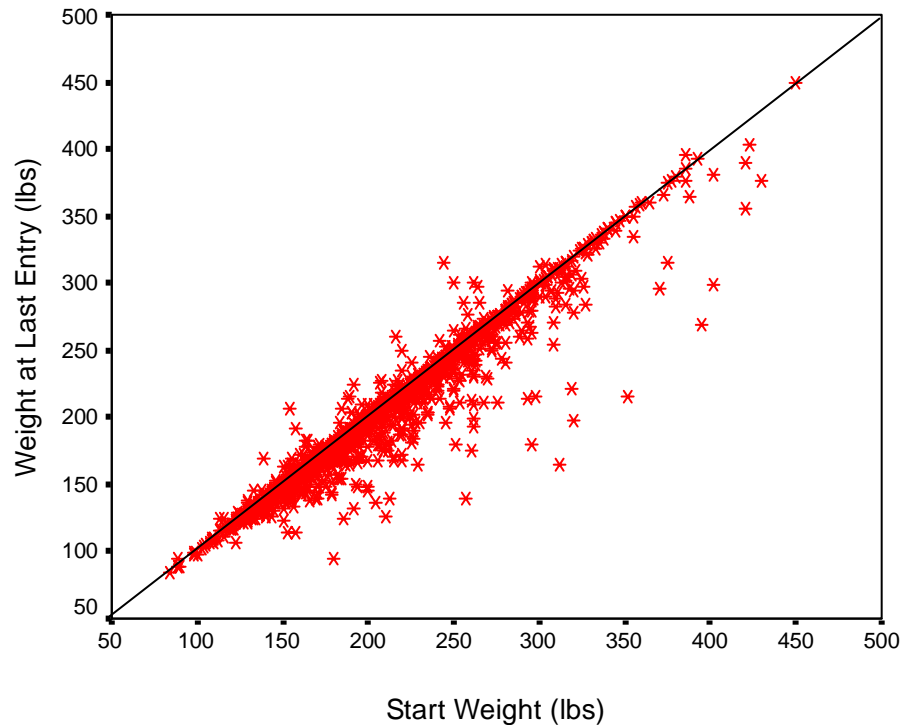
Figure 7: Scatterplot of Starting Weights and Goal Weights



Note. The diagonal line represents values for which starting weight would be equal to goal weight. Any asterisks above the line represent a desired increase in weight. The degree to which asterisks are below the line represent the degree of desired weight loss.

As shown in Figure 8, the degree to which people lost weight over the course of blogging was very small for most bloggers.

Figure 8: Scatterplot of Weight Loss for Bloggers in the Corpus for whom Start Weight and End Weight Information were Available.



Note. The diagonal line represents values for which starting weight would be equal to end weight. Any asterisks above the line represent an increase in weight over the course of blogging. The degree to which asterisks are below the line represent the degree of weight loss achieved over the course of blogging.

SAMPLE

Since the goal of the dissertation was to examine self-change over time, a sample of blogs were selected for analyses. That is, in order to sample from bloggers who had been committed to blogging about weight loss beyond a couple posts, and to sample many words per blog, a sample of blogs were selected for further analysis.

Below are the selection criteria that were considered in creating a sample of blogs for analysis (see Table 4):

- a) Blogs must have posted over the course of a minimum of 15 consecutive weeks since the start of the blog, but not necessarily with an entry each week.
- b) Blogs must have had a minimum of 15 entries within the first 15 weeks. This number was just below the mean number of entries, but ultimately, the 15 entry criteria for the sample was arbitrary,
- c) The sex of the blogger had to be apparent within the blog and bloggers had to be female. Due to the small number of male bloggers (2.96% of corpus), and since males lose weight differently and have different cognitive and emotional responses to eating and dieting (De Souza & Ciclitira, 2005; Economos, Hildebrandt, & Hyatt, 2008; Presnell, Pells, Stout, & Musante, 2008), and tend to use words differently than females (Newman et al., 2008), blogs by males were excluded from the sample.
- d) Obvious repeat attempts by the same blogger were excluded. For example, if a screen name (e.g. BigDieter) appeared that was similar to another screen name in the sample (e.g. BigDieter2), the blog with the later start date was excluded since this was likely the same user. The goal was to have one blog per person, and so even if the blogger stated that their blog was a second attempt, as long as there were no other blogs by that blogger in the sample, the blog was included.
- e) Blogs must have updated weight information for at least 50% of posts, with no more than 12 posts with consecutive missing weights.

- f) A single blogger had recorded only weights with nothing else in the entries, and so this blog was excluded from the analyses.
- g) Bloggers who had met inclusion criteria for 15 entries within a minimum of 15 weeks but who had not blogged during weeks 13, 14, or 15 were excluded from the analyses. These bloggers had blog entries beyond 15 weeks, but no entries during weeks 13, 14, or 15 (see Table 5). Since the degree of weight loss increased over time (i.e. the longer people dieted, the more opportunity they had to lose weight), it would have been unreasonable to compare target variables for weight loss at an ending period of 3 weeks to weight loss at an ending period of 15 weeks. Accordingly, only blogs with an active entry for at least the 12 week period were included.

Table 4: Exclusion Criteria for Sample

Criteria	Number of Blogs Excluded by Criteria	Cumulative Number of Excluded Blogs	Cumulative Number of Blogs in Sample
(a) Minimum 15 weeks of blogging since starting blog	446	2084	446
(b) Minimum of 15 entries during the 15 weeks since starting blogs	173	2257	273
(c) Sex of the blogger must be available; female	20	2277	253
(d) Obvious subsequent attempt by a blogger	3	2280	250
(e) Weight information for at least 50% of posts	17	2297	233
(f) Other (1 blogger with weight info but no words for each entry)	1	2298	232
(g) Must have an entry during at least the 12th week of the sample	46	2344	186
Final Sample	186	2344	186

The duration (a) and number of posts (b) criteria resulted in the collection of 273 blogs. From this sample, (c) 11 blogs were excluded because the sex of the blogger could not be determined, and 9 blogs were written by males, (d), 3 blogs were obvious 2nd

attempts by bloggers already in the sample and (e) 15 blogs did not have updated weight information for at least 50% of their posts, (f) 1 blog included only weight information over the course of the 15 week period, but had no written entries, and (g) 46 blogs did not have an entry during weeks 13, 14, or 15, and so were excluded from the sample, resulting in a total of 186 blogs for the sample.

Table 5: Week of Final Post During the 15 Week Sample Period for Blogs in the Sample

Week of Final Post During the 15 Week Sample Period	Number of Bloggers	Cumulative Number of Blogs Excluded from Sample	Cumulative Number of Blogs Included in the Sample
1	0	0	0
2	0	0	0
3	1	1	0
4	2	3	0
5	4	7	0
6	6	13	0
7	3	16	0
8	3	19	0
9	7	26	0
10	9	35	0
11	11	46	0
12	11	46	11
13	13	46	24
14	27	46	51
15	135	46	186
Total	234	46	186

The final sample characteristics (n = 186 blogs; 9,293 entries) can be seen in Table 6. On average, bloggers had a mean of 49.96 entries ($SD = 25.97$). The median number of posts was 46.00. The average word count per blog 12,562.36 ($SD = 10,165.51$). The average word count per entry was 235.72 ($SD = 127.39$).

Age information was available for 138 bloggers in the final sample. Bloggers averaged 31.02 years old ($SD = 9.38$) at the start of their blogs. All bloggers in the sample were female.

Table 6: Descriptive Information on www.dietdiaries.com Sample (n = 186)

Measure	<i>M</i>	<i>SD</i>	Median	Min	Max
Age in Years	31.02	9.38	29.00	18	65
Blog Entry Rate					
# Entries per blog	49.96	25.97	46.00	15	149
# Days per blog	99.66	7.05	103.00	78	105
# Weeks per blog	14.54	.86	15.00	12	15
Total word count per blog	12,562.36	10,165.51	9,683.50	683	56,869
Word count per entry	235.72	127.92	218.92	34.09	1,225.09
Weight Information					
Start weight in lbs	211.37	54.70	206.00	113	420
End weight in lbs	194.18	49.26	185.25	106	390
Goal weight in lbs	142.64	23.71	140.00	105	270
Goal # lbs to lose	68.68	45.95	60.00	2.80	220.00
Change in weight in lbs	-13.83	13.98	-10.90	-71.50	12.50
Percent body weight loss	6.07	5.74	5.41	28.43	8.96
Social Support					
Total comments received	133.89	136.76	88.00	0	795
Comments received per entry	2.83	2.53	1.38	0	7.66
Total word count of comments received	8,619.87	9,865.58	5,234	0	55,988
Mean word count of comments per entry	174.67	111.34	145.85	0	748.93
#Different commenters	27.03	15.01	25.00	0	72
#Different commenters per entry	2.75	1.31	2.47	0	7.44
#Comments to other blogs	98.00	182.32	17.50	0	1279
#Different blogs commented to	23.20	31.06	10.50	0	148

Note. Means for the sample are based on the information within the first 15 weeks of blogging.

TARGET MEASURES

Percent Body Weight Loss

Percent of body weight loss for each blog over the first 15 weeks of blogging was computed by the following: $-100 * [\text{weight at 15 weeks} - \text{starting weight}] / [\text{starting weight}]$. Positive numbers represented weight loss; negative numbers represented weight gain. Starting weight was typically available on the first entry. However, when there was no weight information in the first entry (n = 4), starting weight was taken from the first recorded weight within the first five entries. Similarly, if the ending weight was unavailable on the last entry within the sampling period (n = 6), ending weight was taken

from the last recorded weight within the last five entries. In order to assess the degree to which starting weight alone was predictive of weight loss, and the degree to which smaller and larger goals were predictive of weight loss, starting weight was examined, as was the discrepancy between starting weight and goal weight (i.e. “goal discrepancy” = starting weight – goal weight).

Note that the amount of weight loss in the 15 week period (13.83 lbs) was much smaller than that found in a meta-analysis of over 700 studies on therapeutic weight loss treatments for obesity. Specifically, the meta-analysis reported a weight loss of about 24.2 lbs across studies for the same time period. There are several possible reasons for why the weight loss in blogs did not meet the mean weight loss reported in the meta-analysis. Bloggers in the current sample started at a mean weight of 211.37 lbs, suggesting that some proportion of the sample would not be considered clinically obese and therefore would not lose as much weight upon starting their diets as a given sample consisting of all obese participants. In addition, bloggers in the current sample were not necessarily receiving therapeutic weight loss treatments.

Blogging Rates

Bloggers started posting about five to six entries per week, but this rate decreased over the course of the 15 week period. There was a slightly lowered rate of blogging (about three entries per week) in the latter half of the 15 week period. Overall, this pattern suggested that bloggers in the sample tended to start off blogging almost every day, but blogging rate decreased steadily each week. By about the eighth week of blogging, bloggers were blogging three or fewer entries per week. For the research questions,

measures of blogging rates were the mean number of entries per week, the number of active weeks, and the word count of entries.

Comments

The number of comments to a blog during the specified time period for the sample (excluding any rejoinders by the blogger) was summed. The number of different people who commented to the blog during the sample time period was also counted. All comments to a blog during the specified time period for the sample (excluding any rejoinders by the blogger) were concatenated and processed using LIWC2007. The categories of interest were positive and negative emotion words.

Finally, the number of comments made by the blogger to other blogs and the number of different blogs commented to during the sample period were also counted as measures of investment and social connectedness by the blogger to the blog community.

Word Counts

Only the words in the entries, excluding the dates, comments, and headings of the entries, were processed using LIWC2007. LIWC2007 reported the percentage of words in each blog that belongs to one of the categories in its internal dictionary. The categories of interest were first person singular pronouns, positive and negative emotion words, cognitive mechanism words, ingestion words, and social words.

In order to assess whether the target LIWC categories were measuring self-focus (using first person singular pronouns), making realizations (using cognitive mechanism words), food diaries (using ingestion words), social support (using social words), and expressions of negative and positive emotions (using negative and positive emotion

words), four independent judges coded a random selection of 116 blog entries, each entry from a different blogger. Three of the judges were female (2 undergraduate dieters, 1 post-graduate non-dieter) and one of the judges was male (post-graduate non-dieter). The judges were between the ages 21 and 25.

The judges were blind to the hypotheses and were asked to rate each entry for the degree to which the blogger was self-focused, had made realizations, felt social support (either online or offline), and was expressing or suppressing negative and positive emotions (see Appendix A). Each blog entry was also rated for the degree to which the entry resembled a food diary (i.e. was a list of the types and amounts of food consumed in a given time period). Raters were given the option of skipping the item if they had “really tried, but felt the item could not be rated for that blog entry” for several of the items, resulting in an uneven number of ratings for each item.

Excluding entries with a word count less than 50 words resulted in 103 entries to compare the average judges’ ratings and the LIWC categories (see Table 7).

Table 7: Inter-Rater Reliabilities of Judge’s Ratings of Entries and Correlations to LIWC Categories

Item	Rater Reliability				Correlations with LIWC Categories $r(103)$					
	ICC	<i>F</i>	<i>df</i>	<i>p</i>	I	Negemo	Posemo	Cogmech	Ingest	Social
Self-focus	.32	1.47	113, 339	.005	.17+	.02	-.30**	.12	-.01	-.34***
Expressing negative emotions	.86	7.13	113, 339	.000	.05	.24*	-.25*	.14	-.07	-.02
Suppressing negative emotions	.59	2.46	106, 318	.000	-.09	.18+	-.01	-.04	-.09	-.09
Expressing positive emotions	.84	6.36	113, 339	.000	-.03	-.15	.48***	-.02	-.03	.31***
Suppressing positive emotions	.48	1.91	107, 321	.000	.10	-.02	.09	.16	.09	-.19*
Making realizations	.66	2.98	103, 309	.000	.15	.14	.24*	.18+	.12	.08
Food diary	.88	8.24	112, 336	.000	-.20*	.02	-.15	-.11	.65***	-.31***
Social Support	.80	4.98	84, 252	.000	-.07	.14	.40***	.06	-.13	.42***

Note. + $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

As can be seen in Table 8, all of the items were rated reliably. That is, judges agreed on the degree to which entries were self-focused, expressing negative emotions, food diaries, etc. The strongest relationship between judges' ratings and the LIWC categories was the degree to which entries were rated to be more like food diaries (i.e. a record of the kinds and amounts of food or calories consumed within any given time period) and LIWC's ingestion word category. This was expected, as this is a content word category. Content word categories are expected to be highly correlated with judges' ratings of a topic, since content word categories include most high base rate words that are on a given topic. Note that ratings of food diaries were strongly correlated with the ingestion word category (e.g., *breakfast, carbs, meat, pasta*; $r(103) = .65, p < .001$) but not with the health word category (e.g. *binge, exercise, heartburn, physical, unhealthy*; $r(103) = .01, p = .913$), demonstrating the discriminant validity of the judges' ratings (see Appendix B).

The degree of social support perceived by judges' was highly correlated with the use of the LIWC social word category, which is also a content word category with some personal pronouns included. For this item, judges were asked to rate the degree to which the blogger of each entry felt social support either online or offline. However, of the correlations of the judges' ratings with the social word subcategories (i.e. friends, family, and humans), the judges' ratings were significantly correlated only with the friends category (e.g. *boyfriend, buddy, fellow, friendship*; $r(103) = .27, p = .006$) and not with the family (e.g. *daughter, husband, married, relatives*; $r(103) = .06, p = .640$) or humans

(e.g. *adult, baby, female, individual*; $r(103) = .07, p = .502$) categories (see Appendix C). Interestingly, the judges' ratings were significantly correlated only with second person pronouns (e.g. *y'all, you, yours, yourself*; $r(103) = .35, p < .001$), and marginally correlated with first person plural pronouns (e.g. *let's, ourselves, us, we*; $r(103) = .17, p = .090$). No other correlations of judges' ratings with personal pronouns were significant. Taken together, these suggest that much of the perceived social support referred to online support (i.e. the audience of the blog, or the diet community).

The degree to which judges rated a blogger as having made realizations showed a marginally significant correlation with the use of LIWC's cognitive mechanism word category. However, this correlation was the least strong of all the content word categories, perhaps because LIWC's cognitive mechanism word category is a superordinate category for many different subcategories: causal words, insight words, discrepancy words, tentative words, certainty words, etc. Of LIWC's cognitive mechanism subcategories, judges' ratings of making realizations correlated strongly with insight (e.g., *discover, realize, understand, etc.*; $r(103) = .38, p < .001$), and causal words (e.g. *affect, cause, why, etc.*; $r(103) = .24, p = .013$; see Appendix D). There were no significant correlations with any other cognitive mechanism subcategories.

The emotion word categories can also be considered content categories, and indeed, these were significantly correlated with the judges' ratings of the expression of particular emotions. Specifically, judges' ratings of the degree to which bloggers were expressing negative and positive emotions were strongly correlated with the rate of negative and positive emotion word use in the entries, respectively. It was unexpected

that the use of negative emotion words was also non-significantly positively associated with the judges' ratings of the degree to which bloggers were suppressing negative emotions. These suggest that the use of negative emotion words is merely a reflection the degree to which bloggers might be experiencing negative emotions, whether expressed or suppressed. Ratings of expressing negative emotions were significantly correlated with two of the three subcategories of negative emotion words (see Appendix E). Specifically, the ratings were significantly correlated with the anxiety (e.g. *anxious, guilt, stress, worry*; $r(103) = .32, p = .001$) and the anger (e.g. *furious, disgust, offensive, temper*; $r(103) = .21, p = .033$) subcategories, but not with the sadness subcategory (e.g. *cry, discouraged, helpless, ruined*, $r(103) = -.12, p = .247$).

Finally, in the only non-content category, judges' ratings of self-focus were only marginally correlated with the use of first person singular pronouns. Judges had been asked to rate the degree to which the bloggers' attention had been focused on themselves or inward. Upon closer inspection of the ratings with the different types of 1st person singular pronouns used, it seemed that this set of instructions was slightly biased in rating possessive instances of the self (see Appendix F). Specifically, regarding types of first person singular use, judges ratings of self-focus were only correlated with the possessive and reflexive forms (e.g., *mine, my, myself*, etc.; $r(103) = .23, p = .021$), but not with the subject form (e.g., *I, I'd, I've*, etc.; $r(103) = .07, p = .491$) nor with the object form (*me*; $r(103) = -.06, p = .545$).

Note that first person singular pronouns were negatively associated with the degree to which the blog entries were rated to be food diaries, and non-significantly

positively associated with the degree to which judges perceived bloggers to be making realizations (see Appendix G). These findings make sense: a food diary isn't very personal, and in order to use many of the cognitive mechanism words, a pronoun is required (e.g. "I knew", "my understanding", "affected me", etc.). From the correlations, it appeared that the strongest cue used to assess self-focus was not first person singular pronouns, but rather a low usage of social words. This might result from the high perceivability and salience of content words. Specifically, since social words are more perceivable and salient, judges may have been using a high rate of social words to mean less self-focus, and a low rate of social words along with an absence of ingestion words to mean more self-focus. For correlations of the judges' ratings with all other LIWC categories, see Appendix H.

SUMMARY

Overall, the descriptives of the entire corpus showed that diet blogs are rich sources of information for examining the process of self-change. The information contained within a blog, because it is a public, time-stamped record of written accounts of daily life, comments, and potential social links within a community affords the possibility for examining natural language, blogging rates, and social support in the self-change process. One thing that was striking upon examining the entire corpus was that many bloggers dropped out shortly after starting their blogs, and most bloggers did not come even remotely close to achieving their goal weights. Indeed, there are many accounts of failures in pursuits at self-change, as they often occur repeatedly before people experience some success at self-change if at all (Herman & Polivy, 2003). That the

records of dropouts were archived on the Internet along with those who became dedicated bloggers provided a staggering picture of how many people fail at self-change efforts.

Even with the sample, evidence of stamina for blogging about weight loss appeared to rapidly decrease within the first couple weeks. The rate of entries per week dropped sharply with the first few weeks, to a much lowered and steadier rate after the first couple months of blogging. The small portion of the sample that met the duration and minimum entry requirements (15 entries in 15 weeks) also pointed towards the difficulty in maintaining a diet blog; many people find it difficult to maintain the same degree of effort in persisting through a diet as they had upon starting the task (Polivy & Herman, 1999). The overwhelming number of failures within the diet blogs makes the question of the determinants of weight loss all the more compelling: what predicts weight loss in blogs?

Chapter 4

Research Question 1: What Predicts Weight Loss in Blogs?

Previous research has shown that breaking one's diet is associated with high self-focus, and suppressed emotions, and that losing weight is associated with high social support, and high self-monitoring, such as keeping regular food diaries. The current study tests whether these features of successful weight loss in lab and clinical studies hold up in blogs devoted to weight loss. Blogs in the sample were examined for language markers of self-focus, self-reflection, and emotionality, along with start and goal weight information, blogging rates, and social support. A linear regression was conducted to determine significant predictors of weight loss in the blogs.

In order to examine potential correlates of weight loss, the sample who had blogged for a minimum of 15 entries for at least 15 weeks described in the previous chapter (n = 186) was examined. All blog entries up to 15 weeks for bloggers in the sample were compiled into a single plain text file for each blog. All comments to the entries during this 15 week period were compiled into a single plain text file for each entry (excluding rejoinders by bloggers). The blogs and comments were processed using LIWC2007 in order to determine the percentages of words devoted to target word categories in the blogs (see Appendix I for usage rates of all LIWC categories). The blogging rate during this time was assessed by counting mean entries per week, along with word counts per entry, and the number of active weeks (i.e. the number of weeks for which there was at least one blog entry). The degree and kinds of social support were

assessed by counts of the comments made and received, along with the number of different bloggers who responded to and received comments from a blogger.

In order to determine the processes associated with weight loss, each of the target variables was correlated with percentage of body weight loss. As can be seen in Table 8, the heavier bloggers were, and the more ambitious the goal weight, the greater the weight loss they achieved over the course of the 15 weeks. Weight loss was associated with the use of first person singular pronouns and positive emotion words in the blogs. Percentage of body weight loss was also significantly associated with having a greater number of weeks in which there was at least one blog entry made. Although the content of the comments was not associated with weight change, the number and connections to the blog community were strongly associated with weight loss. Specifically, receiving many comments, from many different commenters, as well as commenting on other blogs in the community, to many different bloggers, were associated with weight loss. See Appendix J for correlations between the target variables.

Table 8: Correlations between Percent Body Weight Loss and Mean Levels of Target Variables During the 15 Week Period

Dependent Variables	Examples	<i>M</i>	<i>SD</i>	<i>r</i> with %Body Weight Loss	<i>p</i>	<i>df</i>
Start Weight		210.46	55.08	.34	.000	186
Goal Discrepancy		68.16	45.90	.38	.000	178
Self-Focus						
First person singular pronouns	I, me, my	9.25	1.61	.18	.013	186
Emotions						
Negative emotion words	angry, sad, ugh	1.60	.38	.11	.147	186
Positive emotion words	excited, Lol, yay	3.99	1.18	.17	.020	186
Cognitive Mechanisms						
Cognitive mechanism words	caused, realize, understand	15.46	1.51	-.02	.785	186
Self-Monitoring						
Focus on Food						
Ingestion words	binge, carbs, eat	3.06	1.39	-.07	.346	186
Rate of Blogging						
Mean entries per week		3.32	1.73	.04	.603	186
Number of active weeks		12.35	2.88	.19	.011	186
Total word count		12,562.36	10,165.51	.03	.669	186
Social Support						
Social words in blogs	family, share, us	5.34	2.05	-.01	.929	186
Total comments in sample		133.89	136.76	.15	.041	186
#Different commenters per blog		27.03	15.01	.15	.047	186
Total word count of comments received		8,619.87	9,865.58	.12	.118	186
Negative emotion word use in comments received	damn, hate, sucks	1.74	.37	-.05	.530	183
Positive emotion word use in comments received	excited, Lol, yay	6.94	1.27	.11	.156	183
#Comments to other blogs		98.00	182.32	.20	.006	186
#Different blogs commented to		23.20	31.06	.21	.004	186

Note. Positive correlations indicate that the more a person was high on a target variable, the more weight loss. Negative correlations indicate that the more a person was high on a target variable, the more weight gain.

The relationship of body weight loss during the sample time period was correlated with all of the aforementioned dependent variables. One case was excluded in the correlational analyses of LIWC percentages for comments, since the person only had a single comment consisting of 3 words. All other blogs were considered in computing correlations with the other target variables.

For the regression analysis, all of the variables were entered in a stepwise linear regression equation to predict body weight change. A stepwise linear regression equation

was chosen over other methods of regression methods since all target variables were theoretically driven, but the strength of each of the variables in predicting weight loss in blogs was unknown. A stepwise method allowed for an exploratory examination of the strongest predictor variables for weight change. For those cases in which there was missing information (i.e. for those blogs that had no comments), the mean of the target variables replaced these missing cases in order to maximize the amount of information used in the regression model.

In running the original regression analyses, there was 1 blogger who evidenced a predicted standardized residual 4 standard deviations from the mean of percent body weight change. This case was excluded from the regression analyses (see Appendix K for the original regression analyses). In running the stepwise regression with the outlier excluded, the sixth step in the regression produced a larger R^2 and adjusted R^2 value than did the seventh step of removal, which had removed a variable (first person singular pronouns) early in the iterations (third step), suggesting that this was a strong predictor. Therefore, the regression analysis was run using a removal value of $p = .15$, and the final model in this stepwise regression produced a model with the same predictors as in the sixth step of the previous regression that had accounted for the largest proportion of variance (Nau, 2009). The final regression model was significant, $F(6, 178) = 10.35, p < .001, R^2 = .26$. The regression equation predicting percent body weight loss was (see Table 9):

$$\% \text{ Body Weight Loss} = - 4.87 + .04 (\text{goal discrepancy}) + .03 (\text{number of other bloggers commented to}) + .36 (\text{first person singular pronouns}) + .31 (\text{number of active weeks}) - .53 (\text{social words}) + .67 (\text{positive emotions words})$$

Table 9: Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>	<i>R</i> ² Change
Constant	-4.87	2.87		1.70	.091	
Goal discrepancy	.04	.01	.36	5.41	.000	.14
Number of bloggers commented to	.03	.01	.19	2.64	.009	.04
First person singular pronouns	.36	.24	.10	1.50	.136	.02
Number of active weeks	.31	.13	.17	2.34	.020	.02
Social words	-.53	.20	-.20	2.67	.008	.02
Positive emotion words	.67	.34	.14	1.98	.049	.02

Note. $F(6, 178) = 10.35, p < .001, R^2 = .26$. Positive betas indicate that the more a person was high on a dependent variable, the more weight loss. Negative betas indicate that the more a person was high on a dependent variable, the more weight gain.

Goal discrepancy, self-focus, number of active weeks, and social support were all significant predictors of weight loss. While not all measures of blogging rate and social support were found to be significant predictors of weight loss, the number of active weeks on a blog, the number of different bloggers commented to, and positive emotion words were found to be predictors of weight loss. The use of social words in the entries was predictive of weight gain.

Having a larger goal discrepancy (i.e. more pounds to lose from starting weight) was predictive of weight loss. Although unrealistic goals are less likely to be realized than realistic goals, the goal discrepancy measure was not necessarily an unrealistic goal if one considers that to evaluate whether a goal is realistic, one must also consider the speed and ease of change (Herman & Polivy, 2003). It is possible, for example, that having large goals were associated with longer time frames of achieving the goal and with less extreme food restriction and exercise plans, suggesting fewer delusions about the difficulties of self-change, and consequently fewer disappointments along the self-

change process. Therefore, it is possible that having a larger goal was associated with an expectancy of being more dedicated to the blog and the weight loss process.

A relatively high use of first person singular pronouns in the blogs was a significant predictor of weight loss. The higher the degree of self-focus (i.e. the more that bloggers used first person singular pronouns), the more weight bloggers could be expected to lose. Unlike the self-awareness theory of overeating, some degree of self-focus or self-reflection was actually helpful in blogging about weight loss. It cannot be ruled out that dieters refrained from blogging during times of hyper self-awareness, since posting on a public blog on a personally threatening topic may have been expected to exacerbate self-awareness. However, the data suggests that some level of self-focus or self-reflection was good for keeping in line with bloggers' self-change goals, as is important for most psychotherapies.

Positive emotion word use in the blog entries was predictive of weight loss, whereas negative emotion word use in the blog entries was not a significant predictor of weight loss. Note, however, that negative emotion word use was correlated with weight loss in the expected direction. Overall, these support that expressing emotions is good for diets, but here, positive emotional expression was found to have a stronger relationship with weight changes.

Surprisingly, while social word use in the blogs was not correlated with percent body weight loss, it appeared in the regression as a negative predictor of weight loss. This finding was unexpected, and perhaps suggested that talking more about the social network on one's blog was a worse strategy than talking about personal events. Indeed,

instead of writing entries about perceived social support, the findings showed that it was far more valuable to actively seek out social support. That is, the number of different blogs that a blogger had commented to was a significant predictor of weight loss, suggesting that being broadly connected, especially to the blog community was important for weight loss. In the current analysis, it cannot be determined whether making comments to other blogs caused weight loss, or whether losing weight led to making comments to other blogs. Perhaps extraverted, gregarious people were more likely to make comments and to lose weight. A look at the mediating variables, or an analysis on an entry by entry level is required to assess the causal pathway of weight loss.

The number of active weeks (i.e. the number of weeks for which there was a blog entry made) was a significant predictor of weight loss whereas the length of blog entries was not. The number of entries per week was also not a significant predictor of weight loss. Taken together, these suggest that merely checking in every week (and writing in a personal style and expressing emotions) without necessarily having to write a long blog entry or doing so more frequently, was predictive of weight loss.

The use of ingestion words was not predictive of weight loss in the current analysis. This was surprising, given that the use of ingestion words were strongly associated with judges' ratings of entries being food diaries, and keeping food diaries, in turn, have been found to be consistently associated with weight loss. However, food diaries may be a more successful strategy for clinical weight loss programs where nutritionists, counselors, psychologists, and access to other weight loss participants is limited, and where adherence to independent weight loss activities are needed. Given that

the associations between first person singular pronouns and social support with weight loss were so strong, the findings suggest that for blogs, having a personal blog in which personal thoughts and events are shared with other bloggers, and support is actively sought out and/or offered to others, is a successful strategy for weight loss online. Perhaps an online food diary elicits fewer comments than a more emotional and personal account of the day's events, challenges, and victories with the weight loss process. Indeed, these analyses point towards the need for future research to assess what kinds of blog entries are predictive of receiving more comments, of encouraging bloggers to come back every week, and of seeking out other people's blogs. Further social network analyses can be conducted to determine what features lure people to particular blogs, and to identify those commenters who are the largest influencers of weight loss by other bloggers.

Overall, these results point to the importance of personal, narrative-style blogs. Self-monitoring, as measured by ingestion word use, entries per week, and total word count, was not as strong predictors of weight loss as expected. However, investment in the online social network was a significant predictor of weight loss. The findings presented here suggest that future research and treatment programs should emphasize the benefits of a social network in their programs.

Chapter 5:

Research Question 2: Do changes in word use over time lead to weight loss?

Recall that changes in word use over time in written narratives were predictive of improvements in physical health in the months following writing. Specifically, changes in pronoun use, and increases in cognitive mechanism word use have been found to be associated with improved health after expressive writing (Campbell & Pennebaker, 2003; Pennebaker, et al., 1997). Also, a greater use of cognitive mechanism words and having reported contemplating change in retrospective accounts of self-change have been found to be associated with success in attempts at self-change (Sobell et al., 2001). The current investigation sought to examine the process of blogging about dieting as bringing about cognitive, emotional, perspective, and social changes that might influence weight loss. Changes in word use, blogging rates, and social support were correlated with percent body weight loss. The difference scores on all these variables were simultaneously entered into a stepwise regression that also included starting weight and goal discrepancy reported only at the very beginning of the 15 weeks to predict percent body weight loss.

In order to examine changes in blogging that may lead to weight loss, the same sample of bloggers who had blogged for a minimum of 15 entries, and a minimum of 15 weeks were selected. However, for the current analyses, the first 7 weeks of entries were compared with the next 8 weeks of entries. The files were not split evenly (i.e. into 7.5 weeks) since 1) the measurement of rate used weeks as units, 2) there were typically more entries in the first half of blogging than in the latter half of blogging for the sample,

and 3) the final entry in the sample for some bloggers occurred in weeks 12, 13, or 14 (these bloggers had met the minimum 15 weeks requirement, but they happened to not have blogged during weeks 13, 14, and/or 15). The blogs were examined for changes in self-focus, self-reflection, and emotionality, and for blogging habits such as blogging rate and social support.

Table 10 shows how weights, language variables (which includes halves analyses that have total word counts > 50), rates of blogging, and comments differed in the first and second halves of the sample (for descriptive on all other LIWC categories, see Appendix L). Overall, the amount of weight lost in the second half of blogging was less than the amount of weight lost in the 1st half of blogging.

Table 10: Mean Differences on Target Variables in the First Half and the Second Half of Blogging

Dependent Variables	First Half (Weeks 1 through 7)			Second Half (Weeks 8 through 15)			Paired- <i>t</i>	<i>p</i>	<i>df</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>			
Percent Body Weight Loss	186	.04	.04	186	.02	.03	5.77	.000	185
Self-Focus									
First person singular pronouns	185	9.33	1.67	184	9.17	1.90	1.77	.078	184
Emotions									
Negative emotion words	185	1.59	.43	184	1.63	.57	1.28	.201	184
Positive emotion words	185	3.94	1.20	184	4.08	1.35	1.47	.143	184
Cognitive Mechanisms									
Cognitive mechanism words	185	15.35	1.61	184	15.59	1.72	1.55	.123	184
Self-Monitoring									
Focus on Food									
Ingestion words	185	3.30	1.45	184	2.65	1.47	6.43	.000	184
Rate of Blogging									
Mean entries per week	186	4.05	2.02	186	2.70	1.83	11.12	.000	185
Number of active weeks	186	6.16	1.39	186	6.19	2.05	.18	.854	185
Total word count	186	6,991.77	5,328.58	186	5,570.59	5,653.08	4.65	.000	185
Social Support									
Social word use in the blog	185	5.10	2.05	184	5.66	2.34	4.52	.000	185
Total comments in half	186	71.06	67.33	186	62.75	78.55	2.19	.030	185
#Different commenters	186	18.41	10.40	186	16.06	11.99	3.58	.000	185
Total word count of comments received	186	4,586.73	4,856.56	186	4,029.78	5,715.13	1.95	.052	185
Negative emotion word use in comments received	182	1.71	.44	180	1.81	.56	1.89	.060	178
Positive emotion word use in comments received	182	6.92	1.43	180	7.25	2.15	2.02	.045	178
#Comments to other blogs	186	51.83	95.36	186	46.18	99.55	1.12	.266	185
#Different blogs commented to	186	15.44	20.37	186	12.85	21.35	2.68	.008	185

Note. LIWC row numbers indicate percentages of words as a function of the blog half.

Blogging rates decreased over time. Many people are enthusiastic and can stick to goals for short periods of time, but motivation and vigor for starting diets tends to sharply decrease immediately after resolving to diet (Polivy & Herman, 1999). The mean entries per week, and word counts decreased over time. There was no difference in the number of active weeks in the first and second halves of blogging. However, this could be because the second half of blogging was based on a total possible eight weeks, and the first half of blogging was based on a total possible seven weeks.

Overall, bloggers tended to decrease in their use of first person singular pronouns, and ingestion word use from the first to the second half of the 15 weeks of blogging.

There were non-significant trends for bloggers to increase in their use of emotion words, cognitive mechanism words, and social word use. These patterns suggest that bloggers were writing less about their own food diaries, writing more about personal or emotional topics, and were becoming increasingly connected with their online support community.

The mean number of comments made and received decreased from the first to the second half of blogging. In addition, comments tended to include a higher proportion of positive and negative emotion words over time. These suggest that while the overall number of communications with other bloggers may have decreased, the emotionality of communication with a few other bloggers stayed the same or increased, suggesting a smaller but more meaningful or close network.

In order to assess the relationships between weight loss and changes in word use, blogging rate, and social support, difference scores were computed between the 2nd half and 1st half of the blog entries for each of the target variables (see Table 12). The relationship of body weight loss during this time was correlated with the difference scores in the aforementioned dependent variables (see Appendix M for all other LIWC categories).

As can be seen in Table 11, opposite to predictions, weight loss was significantly associated with a decreasing use of cognitive mechanism words. A decrease in first person singular use showed a non-significant trend to be associated with weight loss. No predictions had been made regarding changes in emotion words or ingestion words being

related to weight change. There were no significant relationships between these word categories and weight change.

Rate of blogging was strongly associated with weight loss. Specifically, an increase in the number of active weeks of blogging from the first to the second half of blogging was significantly associated with weight loss. Changes in word count were not significantly associated with changes in weight. Note that blogging many entries per week was actually associated with weight gains, suggesting that merely “checking in” on the blog community by writing a short update and commenting to others’ blogs was a good strategy for weight loss. Indeed, instead of spending copious amounts of time (or rather, words) on one’s own blog, and writing about one’s self (recall that an increase in first person singular pronoun use was non-significantly associated with weight gain), increasing attention on eliciting comments from the social network seemed to be a better strategy for weight loss.

In terms of social support, there was an overall decrease in the amount of comments received and made to other bloggers in the second half of blogging relative to the first half of blogging. There was also an overall decrease in the number of different commenters to a blog and a decrease in the number of different blogs that a blogger commented to. The less of a decrease in the number of comments received and in the number of different commenters, the more weight loss bloggers experienced. Similarly, the less of a decrease in the number of comments made to other bloggers, and in the number of different bloggers one made comments to, the more weight loss bloggers experienced. However, changes in the use of social words in the blogs, and changes in

word use within the comments were not significantly associated with percent body weight gain. See Appendix N for correlations of the target variables with each other.

Table 11: Correlations between Percent Body Weight Loss and Difference Scores between First and Second Halves of Blogging during 15 Weeks of Blogging

Variables	N	Mean Difference (Second Half – First Half)		r with %Body Weight Loss	p	df
		M	SD			
Self-Focus						
First person singular pronouns	183	-2.09	1.55	-.12	.122	183
Emotions						
Negative emotion words	183	.07	.71	-.02	.765	183
Positive emotion words	183	.10	.84	-.02	.747	183
Cognitive Mechanisms						
Cognitive mechanism words	183	.17	1.53	-.16	.033	183
Self-Monitoring						
Focus on Food						
Ingestion words	183	-.60	1.26	-.05	.546	183
Rate of Blogging						
Mean entries per week	186	1.24	1.11	-.19	.009	186
Number of active weeks	186	.03	1.99	.21	.004	186
Total word count	186	-1,421.18	4,167.07	.10	.173	186
Social Support						
Social word use in the blogs	183	.54	1.50	.04	.570	183
Total comments in sample	186	-8.32	51.80	.18	.014	186
#Different commenters	186	-2.34	8.94	.18	.012	186
Total word count of comments	186	556.95	3,887.41	.16	.035	186
Negative emotion word use in comments received	179	.09	.66	.02	.841	179
Positive emotion word use in comments received	179	.33	2.16	-.01	.948	179
#Comments to other blogs	186	-5.65	69.03	.17	.019	186
#Different blogs commented to	186	-2.58	13.15	.21	.004	186

Note. Positive correlations indicate that the more a person increased on a dependent variable in the second half of blogging, the more weight loss over the 15 week period. Negative correlations indicate that the more a person increased on a dependent variable in the second half of blogging, the more weight gain over the 15 week period.

All variables, including the start weight and goal discrepancy at the beginning of the 15 week period assessed in Research Question 1 were entered into an exploratory stepwise linear regression to predict percentage of body weight gain over the course of

the 15 weeks. Any missing variables for bloggers (i.e. for 1 blogger who had no words in the first half of the sample, and for those cases in which there were no comments (5 blogs in the first half, and 6 blogs in the second half) were replaced with the mean in the regression analyses in order to maximize the use of the remaining information for those bloggers. In order to use all cases in the regression, any percentages of word counts based on less than 50 words were deleted and replaced with the sample mean value.

The removal of one outlier in the original regression equation (see Appendix O) resulted in a significant regression equation, $F(3, 181) = 16.57, p < .001$, with the same predictor variables for percent body weight loss (see Table 12). The regression equation accounted for 21.55% of the variance. The prediction equation was:

$$\text{Percentage body weight loss} = .03 + .0004 (\text{goal discrepancy}) + .0006 (\text{difference in comments received}) - .01 (\text{difference in cognitive mechanism words})$$

Table 12: Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>	<i>R</i> ² Change
Constant	.03	.01		4.70	.000	
Goal discrepancy	.0004	.00	.39	5.85	.000	.14
Difference in comments received	.0002	.00	.22	3.26	.001	.05
Difference in cognitive mechanism words	-.01	.00	-.17	2.65	.009	.03

Note. $F(3, 181) = 16.57, p < .001, R^2 = .22$. For goal discrepancy, the positive beta simply indicates that the larger the discrepancy between starting weight and goal weight, the more weight loss. For all other variables, positive betas indicate that the more a blogger increased in blogging from the first to the second half of blogging on that target variable, the greater the weight loss. Negative betas indicate that the more a blogger decreased in blogging from the first to the second half of blogging, the greater the weight loss.

In terms of weight loss, larger discrepancies in goal weight from starting weight were predictive of weight loss. The more that bloggers were connected to the online community through receiving comments, the more weight they lost. On the other hand,

increases in cognitive mechanism words were predictive of weight gains. This latter finding was opposite to predictions. Specifically, it was expected that increases in cognitive mechanism word use would lead to improved mental and physical health, as has been found in the expressive writing literature and in narratives of successful self-change. However, in the current study, the greater use of cognitive mechanism words were likely used when dieters were making excuses or finding reasons for having broken their diet or for ruminating about their weight dissatisfaction.

The difference between previous studies that have found salutary effects by an increase in cognitive mechanism word use and the current study is that previous studies had participants recall traumas or successes that had happened in the past, while the blogs likely represented narratives from the same day or week in which the event was being recalled. Perhaps some temporal distance from an event is required in order to a high use of cognitive mechanism word use to lead to improved health. Indeed, one review on expressive writing has suggested that waiting some time (at least 1 to 2 months after a major upheaval) before expressive writing might improve its efficacy (Pennebaker & Chung, 2007). And while cognitive mechanism word use can indicate flexibility in thinking while forming a new narrative of a negative event over three to four days of writing (Pennebaker et al., 2007), perhaps resolve to lose weight is better for weight loss than is flexibility in goal pursuits.

Perhaps the most interesting pattern of effects in this chapter was from the social support measures. Social word use significantly increased over time. The total number of comments decreased over time, as did the number of different commenters. However,

note (from Appendix M) that the mean number of comments per entry increased over time, suggesting that only some entries were commented to, but that when they were, they received many comments, or were more like conversations between the blogger and commenter, with a few commenters commenting multiple times to an entry. While the number of comments made to other blogs did not change over time, the number of different blogs commented to decreased over time. Taken together, these findings suggest that communication to community members became more specific; bloggers were culling communication to some members while cultivating more frequent communications with a smaller set of bloggers. In terms of the content of comments, both positive and negative emotion word use increased in the comments over time, suggesting that the comments themselves were more personal or emotional. Overall, this snapshot of the social dynamics within the blog community implied that bloggers tended to develop emotional relationships with a small group of bloggers, instead of randomly commenting to as many bloggers as possible.

Chapter 6:

Research Question 3: Can we predict dropouts and weight loss from the first two entries?

People who engage in self-change efforts often fail, and yet they start again and again. This pattern of taking up an old goal despite repeated failures has been termed the false-hope syndrome (Polivy & Herman, 2002). Given that there are so many false starts in the self-change process, and given that there were so many false starters in the blog community, these were worth examining. The first comparison was between dropouts (i.e. those who had blogged between two and five entries) and dedicated bloggers (in the 15 week sample). This comparison addressed whether there were effective start strategies that led to continued blogging about weight loss. Each of the target measures were computed based on the first two entries of each member of these groups. A regression analysis was conducted on the measures to predict group membership (i.e. dropout or dedicated blogger).

The second comparison addressed whether or not the first two entries were able to predict weight loss in the dedicated blogger sample. Each of the target measures was computed based on the first two entries of each member of these groups. A stepwise linear regression analysis was conducted on the measures to predict percent body weight loss over the 15 weeks of blogging.

DROPOUTS VS. DEDICATED BLOGGERS

Dropouts

Dropouts were defined as those who blogged between two and five times (32.13% of the sample). A minimum of two entries was chosen in attempts to sample at least 50 words per blogger. All bloggers who had at least two entries constituted 69.89% of the blog community (n = 1768). Within this 1768 set of blogs, those who had blogged between two to five times were selected (n = 813). However, excluding those whose sex could not be determined, males, those who had not posted their weight for the first two entries, and obvious repeat attempters, the final sample for the dropouts included 445 dropouts.

Dedicated Bloggers

Dedicated bloggers were those who met the criteria for the 15 week sample in the previous chapters (n = 186). Recall that some weights on the first and/or second entries for the 15 week sample had been filled in by using the earliest weight information available for the sample (n = 10). These cases were excluded for the following analyses in order to examine only those cases for which weight differences between the first two entries could be computed.

The final sample for the analyses was composed of 621 bloggers. Specifically, the sample was made up of 445 dropouts, and 176 dedicated bloggers. As can be seen in Table 13, dropouts did not dropout because they had reached their weight loss goals. On the contrary, both groups had about the same weight loss goals (to lose an average of about 60 to 70 lbs when starting the blog), and both groups had about the same amount to lose by their second entry.

Table 13: Mean Age and Weight Differences between Dropouts and Dedicated Bloggers on their Blogging

Dependent Variables	Dropouts			Dedicated Bloggers			<i>t</i>	<i>p</i>	<i>df</i>	Cohen's <i>d</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>				
Age	282	30.62	9.41	129	31.05	9.48	.44	.663	409	-.05
Weight Information										
Start weight (lbs)	445	202.27	51.31	176	211.09	54.93	1.87	.059	618	.15
Goal weight (lbs)	370	139.68	21.23	168	143.01	23.86	1.62	.106	536	.14
Weight to lose (lbs)	370	61.65	42.90	168	68.44	46.32	1.63	.104	535	.14
2nd entry wt discrep from goal wt	370	61.10	42.81	168	67.57	45.86	1.59	.113	536	.14
% body weight loss (2 entries)	445	.29	1.74	176	.37	1.42	.55	.584	619	.04

Note. Positive numbers in % body weight loss represent weight loss. Negative numbers in % body weight loss represent weight gain. Since ages and goal weights were manually assessed from reading the entries, the sample sizes are different for each variable.

Each group was examined for language markers (using only those cases with word counts greater than 50) of self-focus, self-reflection, and emotionality, and for blogging rates, and social support in their first two blog entries only. Counts of their comments were also conducted, along with an examination of positive and negative emotion word use in the comments (using only those cases with word counts greater than 50) from the dates spanning only the first two blog entries (for all other LIWC categories, see Appendix P).

As can be seen in Table 14, dropouts and dedicated bloggers differed in the degree to which they expressed emotions in the first posts. Specifically, dropouts used significantly more negative emotion words, first person singular pronouns, and ingestion words than did dedicated bloggers. They did not differ in the degree to which they used positive emotion words or cognitive mechanism words. These suggest that dropouts were writing about highly negative personal topics or keeping food diaries, which seemed to be poor start strategies. Indeed dropouts tended to use fewer social words, and made much less of an effort to comment to other blogs. Dedicated bloggers, on the other hand, started

out writing significantly longer entries than did the dropouts. In terms of social support, dedicated bloggers used more social words in their entries, and they also received significantly longer and less negative comments. Note that dedicated bloggers commented much more to other bloggers and to many more bloggers than did dropouts. See Appendix Q and Appendix R for correlations between target variables within the first two entries by dropouts and by dedicated bloggers, respectively.

Table 14: Mean Differences between Dropouts and Dedicated Bloggers on Target Variables for the First Two Entries

Dependent Variables	Dropouts			Dedicated Bloggers			<i>t</i>	<i>p</i>	<i>df</i>	Cohen's <i>d</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>				
Starting weight	445	202.27	51.31	176	211.09	54.93	1.87	.059	618	.15
Goal Discrepancy	370	61.69	42.92	168	68.48	46.29	1.66	.098	536	.14
Self-Focus										
First person singular pronouns	431	10.83	2.72	174	10.36	2.29	2.02	.043	603	.16
Emotions										
Negative emotion words	431	1.94	1.13	174	1.69	.81	2.66	.008	603	.22
Positive emotion words	431	3.82	1.72	174	3.68	1.65	.95	.342	603	.08
Cognitive Mechanisms										
Cognitive mechanism words	431	15.53	3.41	174	15.55	2.63	.05	.961	603	.00
Self-Monitoring										
Focus on Food										
Ingestion words	431	4.28	2.72	174	3.94	2.01	1.96	.050	603	.16
Rate of Blogging										
Latency to 2nd entry (days)	445	4.66	16.90	176	2.71	7.43	1.47	.141	619	.12
Total word count	445	364.68	291.99	176	508.03	436.73	4.76	.001	619	.38
Social Support										
Social word use in blog entry	431	3.99	2.57	174	4.46	2.62	2.05	.041	603	.17
Total comments in 2 entries	445	4.24	3.19	176	4.70	3.40	1.57	.117	619	.13
#Different commenters	445	3.98	6.02	176	3.90	2.92	.16	.872	619	.01
Word count of comments	445	364.38	291.99	176	508.03	436.73	2.44	.015	566	.21
Negative emotion word use in comments received	358	1.61	.97	146	1.39	.86	2.44	.015	502	.22
Positive emotion word use in comments received	358	7.80	3.19	146	7.94	3.15	.46	.649	502	.04
#Comments to other blogs	445	.59	1.45	176	2.06	4.38	6.28	.000	619	.50
#Different blogs commented to	445	.55	1.32	176	1.77	3.52	6.29	.000	619	.50

Note. All target variables were based on only the first two entries. Comments received and made to other blogs were counted if they fell on or between the span of the dates of the first and second entry.

In order to determine if group membership could be predicted based on the target variables, a binary logistic regression equation was conducted. A backward stepwise removal method was selected using likelihood ratios in order to find the best model for the data. A backward stepwise removal procedure fits the model to the data, and removes the variable that influences the model the least, until all variables that do not make a significant contribution to the prediction model are removed. The backwards procedure is recommended over the forward procedure to decrease the chances of Type II errors (Field, 2000).

A casewise residuals analysis showed 7 cases with studentized residual values greater than 2.00. These cases were removed. A zero value represented membership in the dropout group, and a value of 1 represented membership in the dedicated group (see Appendix S for correlations of the target variables with group membership). The overall model was significant $\chi^2(5) = 69.49, p = .000$. Table 15 shows the contributions of each variable to the model. Note that the R-statistic represents the partial correlation between the outcome variable and each of the predictor variables; variables with higher numbers contribute more to the overall model than do lower numbers. The valence of the R-statistic indicates the likelihood of the outcome occurring as the predictor variable increases (Field, 2000).

The RL^2 for the final regression equation was computed by dividing the final model χ^2 by the original $-2LL = 69.49/545.02 = .13$. This value can range from 0 to 1, and indicates how much better the model was at predicting group membership with the removal of the other variables.

Table 15: Beta Weights and their Statistical Significance in a Regression Equation Predicting Group Membership

Dependent Variables	Beta	SE	Wald	df	p	95% CI for Exp(β)			R
						Exp(β)	Lower	Upper	
Constant	-.07	.60	.01	1	.909	.93			
Cognitive mechanism words	-.07	.04	3.63	1	.057	.92	.862	1.00	.05
Latency to 2nd entry	-.10	.04	6.61	1	.010	.90	.84	.98	.09
Word count	.002	.00	19.25	1	.000	1.00	1.00	1.00	.18
Negemo in comments	-.27	.13	4.49	1	.034	.77	.60	.98	.07
#Comments to other blogs	.25	.06	19.25	1	.000	1.28	1.15	1.44	.18

Note. $RL^2 = .13$ (Hosmer & Lemeshow), $.15$ Cox & Snell, $.21$ (Nagelkerke). Model $\chi^2(5) = 69.49$, $p = .000$. Dropouts were coded as 0, and dedicated bloggers were coded as 1. Negative betas indicate a greater likelihood of being dropout. Positive betas indicate a greater likelihood of being a dedicated blogger.

The results of the binary logistic regression show that all else being equal, a 1% increase in cognitive mechanism words in the first two entries would decrease the odds of being a dedicated blogger by a factor of .92. For every day that passes between the first and the second entry, there would be a decrease in the odds of being a dedicated blogger by a factor of .90. For every 1% increase in negative emotion words received in the comments, there would be a .77 decrease in the odds of being a dedicated blogger. In order to increase the odds of becoming a dedicated blogger, one would have to have a high word count in the first couple entries. One of the best ways to increase the odds of becoming a dedicated blogger would be to comment more on other blogs. Specifically, with all else being equal, for every comment made to another blog during the first couple entries, the odds of being a dedicated blogger would increase by a factor of 1.28.

The model was able to correctly classify 72.52% of the cases (see Table 16). Although the results of the regression model are consistent with the findings in previous chapters, note that the model does not do much better at predicting group membership

than by simply knowing that dropouts made up a significantly larger proportion of the blog community.

Table 16. Classification Table for Regression Model

	Predicted		%Correct
	Dropout	Dedicated Blogger	
Dropout	277	16	94.54
Dedicated Blogger	103	37	26.43
Overall			72.52

During the first two entries of a blog, using fewer cognitive mechanism words, returning to post a second entry shortly after the first, and receiving few negative emotion words in the comments increased the chances of being dedicated blogger instead of a dropout. Writing a lot within the first couple entries, as well as making lots of comments to the community both increased the odds of being a dedicated blogger. All of these suggest the level of investment of a blogger is evident within the first two posts, and is indicated by writing a lot, coming back early, and commenting to other bloggers.

As discussed in the previous chapter, the use of cognitive mechanism words has been interpreted as an index of flexible thinking, and perhaps in blogging, represents less of a resolve to blog away the pounds. Indeed, self-change progress has been described in a series of stages: precontemplation, contemplation, preparation, action, and maintenance (Prochaska, DiClemente, & Norcross, 1992; Prochaska, Norcross, & DiClemente, 1995). Perhaps those who were using a lot of cognitive mechanism words may still have been in the precontemplation or contemplation stages of self-change. People who move past these stages, and especially those in the action stage of self-change typically achieve higher rates of success in their behavior change.

As found in previous studies of participation on online support forums, negative comments significantly reduced the participation rate on the blog. Specifically, receiving comments with a high use of negative emotion words within the first couple entries was predictive of dropout. One study that examined mood changes in response to making a new commitment to weight loss found that for veteran dieters, making a new commitment to weight loss was associated with increased feelings of depression and increased feelings of hopefulness (Polivy & Herman, 1999). Perhaps the negative emotion word use in the comments was in response to the elevated rates of negative emotion words in the blog entries by dropouts. In any case, negative emotion word use in the comments received from the online community, whether for commiserating or for criticism, was not conducive to later weight loss.

These reiterate several issues brought up in previous chapters for further study. For example, an important feature in understanding whether or not someone stays on a blog, is the degree of positive and negative emotion words found in the comments. These suggest that examining the content of blogs for what elicits the a higher number of comments, and the use of positive and negative emotion word use in comments is an important clue to discovering what leads to weight loss and dedication in blogging about weight loss. Finally, it seems that the investment in a blog community (i.e. the degree to which a blogger makes many comments to the community) is an important determinant of weight loss. This is consistent with previous research that has found that success within an online community for self-help depends on the degree of investment by users (Madara, Kalafat, & Miller, 1988).

PREDICTING PERCENT BODY WEIGHT LOSS IN THE DEDICATED BLOGGERS GROUP BASED ON THE FIRST TWO ENTRIES

As the first two entries were predictive of dropping out, it raised the question of whether weight loss success could be predicted among the dedicated bloggers. As described in Research Question 1, the use of first person singular pronouns, making comments to many different bloggers, and positive emotion word use were correlates of weight loss. Could these or other word categories that were used in the first blog entries predict weight loss?

To test this idea, all target variables were assessed within the first two entries for each blogger, and entered into a regression equation along with start weight and goal discrepancy to predict percent body weight loss achieved during the 15 weeks of blogging. Dedicated bloggers were those who met the criteria for the 15 week sample in the previous chapters ($n = 186$). Recall that some weights on the first and/or second entries for the 15 week sample had been filled in by using the earliest weight information available for the sample ($n = 10$). These cases were excluded for the following analyses in order to examine only those cases for which weight differences between the first two entries could be computed. The final sample for the analyses was composed of 176 bloggers.

The two entries for each blogger were assessed for language markers (for word counts greater than 50) of self-focus, self-reflection, and emotionality, and for blogging rates such as latency to the second entry made and word count. Counts of their social

responses within the period of the first two entries were also conducted (for word counts greater than 50 for the language variables).

As can be seen in Table 18, those with greater starting weights and those with larger goal discrepancies from starting weight were more likely to lose weight over the 15 weeks of blogging. These relationships were the same as those found in previous research questions that assessed the same variables, with this chapter including 10 fewer bloggers. Latency to the second entry was associated with weight loss. Although waiting a longer time to blog a second entry increased the odds of being a dropout, it seems that it was also associated with a chance of losing more weight in the case that the blogger did not dropout of the blog. Using more first person singular pronouns and social words in the first two entries was marginally associated with weight loss, which is consistent with findings from previous chapters that writing in a more personal style and perceiving social support is associated with weight loss (or *anticipating* social support, since these were measurements from the first two entries). Note, however, that overall there were few significant relationships to weight loss based on the first two entries, suggesting that rather weak predictions of weight loss might be gleaned from sampling just the first two entries. See Appendix T for correlations of the target variables within the first two entries with all other LIWC variables, and Appendix R for correlations between the target variables in the first two entries.

Table 18: Correlations between Percent Body Weight Loss and Mean Levels of Target Variables during the First Two Entries

Dependent Variables	Examples	<i>M</i>	<i>SD</i>	<i>r</i> with %Body Weight Loss	<i>p</i>	<i>df</i>
Start Weight		211.09	54.93	.36	.000	176
Goal Discrepancy		68.48	46.29	.40	.000	168
Self-Focus						
First person singular pronouns	I, me, my	10.36	2.29	.13	.087	174
Emotions						
Negative emotion words	angry, sad, ugh	1.69	.81	.08	.276	174
Positive emotion words	excited, Lol, yay	3.68	1.65	.11	.165	174
Cognitive Mechanisms						
Cognitive mechanism words	caused, realize, understand	15.55	2.63			174
				-.05	.165	
Self-Monitoring						
Focus on Food						
Ingestion words	binge, carbs, eat	3.84	2.01	.07	.376	174
Rate of Blogging						
Latency to second entry		2.71	7.43	.16	.032	176
Total word count		508.03	436.73	-.02	.777	176
Social Support						
Social words in blogs	family, share, us	4.46	2.62	-.13	.093	174
Total comments in sample		4.70	3.41	-.04	.578	176
#Different commenters per blog		3.90	2.92	-.04	.640	176
Total word count of comments		333.56	346.62	-.02	.830	176
Negative emotion word use in comments received	damn, hate, sucks	1.39	.86			146
Positive emotion word use in comments received	excited, Lol, yay	7.94	3.15	-.01	.912	146
#Comments to other blogs		2.06	4.38	-.08	.304	176
#Different blogs commented to		1.77	3.52	-.06	.445	176

Note. Positive correlations indicate that the more a person was high on a target variable, the more weight loss. Negative correlations indicate that the more a person was high on a target variable, the more weight gain.

In order to determine if weight loss over the 15 week period could be predicted based on the target variables measured only within the first two entries for each blogger, a stepwise linear regression was conducted. In running the original regression, there was one case with a standardized predicted residual greater than 4 standard deviations. When this case was excluded from the regression analyses (see Appendix U for the original regression analyses), the regression equation was still significant $F(2, 172) = 18.98, p =$

.000, $R^2 = .18$. The regression equation predicting percent body weight loss was (see Table 19):

$$\text{Percent Body Weight Loss} = 4.36 + .05 (\text{goal discrepancy}) - .38 (\text{social})$$

Table 19. Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>	R^2 Change
Constant	4.36	.93		4.69	.000	
Goal discrepancy	.05	.01	.40	5.74	.000	.15
Social words	.38	.15	-.18	2.59	.010	.03

Note. $F(2, 172) = 18.98$, $p = .000$, $R^2 = .18$. Positive betas indicate that the more a person was high on a dependent variable, the more weight loss.

Overall, having larger goals and perceiving social support (or again, anticipating social support within the first two entries) were predictive of the degree of weight loss achieved approximately 15 weeks later. As suggested by the weak correlations between the target variables and percent body weight loss, the regression model based on the first two entries included fewer predictors relative to when target variables were assessed over the entire 15 week period.

SUMMARY

Goal discrepancy was an important determinant of the degree of weight loss achieved over the 15 weeks, although it did not distinguish between dropouts and dedicated bloggers. These findings suggest that goal discrepancy is not an indication of continuing to blog. Rather, goal discrepancy might reflect something about the goal pursuit itself. The use of social words in the first two entries was a positive predictor of weight loss, unlike the finding in Research Question 1, where social word use over the 15 weeks of blogging was predictive of weight gains. These suggest that perceived or anticipated social support was a good strategy for early blogging about weight loss. Note

that social word use in the first couple blog entries was highly correlated with receiving more comments, from a greater number of commenters, and lengthier comments (see Appendix R).

Previous research on New Year's resolutions related to behavioral change found no effects of social support or behavioral skills on sticking to a resolution (Norcross, et al; 1989; Norcross & Vangarelli, 1989). Since the current investigation found large effects for social support, perhaps people who blog about dieting are more likely to welcome a public audience or at least receive support from other dieters, and therefore social support effects for change may have been stronger here than in reports of self-change offline. Perhaps writing about personal events led to making social connections more easily than did writing about non-personal events, but when these personal narratives elicited negative comments, bloggers dropped out.

An analysis of the first two entries was able to predict whether or not a blogger would dropout or blog for at least 15 weeks. An analysis of the first two entries was also moderately able to predict the degree of weight loss that would be achieved over the 15 weeks. These findings suggest that the first two entries of a given blog could be analyzed for the degree to which bloggers are at a stage in the self-change process (e.g. precontemplation, contemplation, or action stages: Prochaska et al., 1992; 1995) that is likely to lead to continued blogging about weight loss, and hint at whether the start strategy could be modified to increase the chances of weight loss. The models for predicting the odds of being a dropout and for predicting the degree of future weight loss based on the first two entries were not strong predictive models. However, since the

results are generally consistent with findings from previous chapters, they suggest that some information could be gleaned for the assessment of continued blogging about weight loss and weight loss success, or for the kinds of assessments that might be made upon intake for behavioral treatment programs.

Chapter 7: General Discussion

Thousands of people use blogs to track their efforts to improve themselves. In addition to blogging sites devoted to reducing alcohol and drug abuse, smoking cessation, controlling spending, sexual, and other behaviors, a number of sites have become popular among people who seek to lose weight. Diet blogs are essentially an archive of the self-change process in everyday life, and so they offer a naturalistic opportunity to track thousands of dieters for potentially long periods of time. The aim of the dissertation was to determine what predicts weight loss in a community of diet blogs. By assessing word use using computerized word count tools, along with blogging rates, start and goal discrepancy information, and social support, several determinants of weight loss were identified. In addition, features of self-change were discovered using naturalistic data recorded throughout the process, instead of through retrospective reports.

The features that had been identified in previous research on successful and failed self-change were assessed in the blogs using computerized word count tools. Previous examinations of retrospective reports of self-change found that successful change stories mentioned contemplating change and having lots of social support. Social support and self-monitoring through high participation rates and keeping frequent food diaries in behavioral treatment programs have also been associated with weight loss in behavioral treatment programs. In experimental lab studies, instances of high negative self-focus have repeatedly been shown to lead to overeating and breaking diets.

Given these findings, a series of hypotheses were made about weight loss in the blogs. For example, self-focus was assessed through the use of first person singular

pronouns and negative emotions were assessed through the use of negative emotion words, and so a high rate of first person singular pronouns and negative emotion words were hypothesized to be associated with weight gain. Food diaries were assessed through the use of ingestion words, and so a high rate of ingestion word use was predicted to be associated with weight loss. Social support was assessed from the number of comments made and received, along with the different number of people who made and received comments from a blogger. These were expected to be positively associated with weight loss. To assess the contemplation of change, the use of cognitive mechanism words were hypothesized to be associated with greater weight loss. In addition, since previous research has shown that the use of cognitive mechanism words tends to increase in healthy narratives about past events, an increasing use of cognitive mechanism words over the course of blogging was hypothesized to be associated with weight loss. Participation rates were assessed through the number of entries made per week, the length of the blog entries, and the number of active weeks on the blog (i.e. the number of weeks for which at least one blog entry had been made). These were hypothesized to be positively associated with weight loss. No firm predictions were made about how starting weight and goal weight discrepancy from starting weight would affect weight loss.

Three research questions were asked that addressed what predicts weight loss, based on assessments of language use, blogging rates, social support, along with starting and goal weights.

The first research question sought to examine what predicts weight loss in blogs. Weight loss was predicted by larger goals, high rates of self-focus, positive emotions in

the blog entries, and participation rates, but not by food diaries. Specifically, using more first person singular pronouns and having a greater number of weeks for which there was at least one blog entry were predictive of weight loss. But the use of ingestion words was not predictive of weight loss. The results suggested that regularly writing about personal events was a better strategy in blogging about weight loss than was keeping an online food diary. In addition, actively seeking out social support was important; the number of different blogs that a blogger commented to was associated with weight loss. Bloggers may have been commenting more to other dieters, leading to weight loss, or perhaps dropping a few pounds led to increased confidence to comment to other bloggers. However, the causal direction of weight loss and social support could not be determined from the analyses. Overall, for weight loss within the blogs, a better strategy for weight loss was to build social connections more than keeping food diaries.

The second research question examined whether changes in blogging were predictive of weight loss. Changes in word use, blogging rates, and social support were assessed during the first one to seven weeks of blogging, and in the next eight to fifteen weeks of blogging during the 15 week sampling period. Weight loss was predicted by larger goals, along with increased participation rates and social support from the first to the second half of blogging. Specifically, increases in the number of active weeks and in the number of comments received per entry over the weeks of blogging were associated with weight loss.

Previous research on expressive writing has shown that an increase in cognitive mechanism words is associated with a multitude of physical and mental health benefits

(Pennebaker et al., 1997). This effect has been interpreted to mean that people who use an increasing number of cognitive mechanism words in their writing about past negative events are actively making meaning of their traumatic or stressful experience, and reorganizing the way they think the world works in order to better cope with the event. Surprisingly, increases in cognitive mechanism words were associated with weight gain. The results were inconsistent with the prediction that an increase in processing negative events would lead to benefits in the weight loss process.

Retrospective accounts of successful self-change have been found to include many instances of cognitive words (Sobell et al., 2001). However, the primary difference between this study and the previous studies that have found salutary effects for cognitive mechanism words is that the previous studies asked participants to recall past events, whereas the current study dealt with current or ongoing topics in a blogger's life. Perhaps some distance from an event is necessary in order for cognitive mechanism words to be associated with positive outcomes. Or, as suggested by the stage theory of self-change (Prochaska et al., 1992; 1995), success is more likely when people have moved beyond the precontemplation and contemplation stages of self-change. The findings suggest that staying in an active search for meaning about why one should diet or about why negative events occurred is not beneficial in the weight loss process.

Previous longitudinal research on people who had made a New Year's resolution about a behavioral change showed that over a fifth had quit their resolutions by January 7th (Norcross et al., 1989; Norcross & Vangarelli, 1989). Indeed, an overwhelming proportion of dieters quit blogging about weight loss within the first few entries of

blogging. Since there was such a high rate of early dropouts (i.e. quitting blogging about weight loss), the third research question examined what could be learned from the first two entries. By examining language use, blogging rates, and social support within the first two entries, perhaps the mindset of dieters entering the blogging world or the social feedback received early on could uncover why dieters dropped out so quickly or why only some had succeeded in losing weight.

First, an analysis was conducted on the first two entries to predict whether or not a blogger would dropout (i.e. only blog between two and five entries) or be a dedicated blogger (i.e. blog at least 15 weeks; the sample examined for the previous research questions). Using a high rate of cognitive mechanism words, receiving comments with a high proportion of negative emotion words, and waiting a longer time to post the second entry since the first entry increased the odds of being a dropout. Blogging long entries in the first couple posts, and making many comments to other blogs increased the odds of being a dedicated blogger. Again, the level of investment put into the blog community was a strong predictor of continuing with the blog. On the other hand, the use of cognitive mechanism words was again found to be predictive of negative outcomes, providing further evidence that contemplation about change or blogging about realizing desires to change was not a steadfast way to start a blog.

Another analysis was conducted to determine whether the first two entries could predict the degree of weight loss that would be achieved by each of the bloggers in the dedicated bloggers group. Goal discrepancy, that is, having larger goals, and perceiving a lot of social support within the first two entries were predictive of the degree of weight

loss achieved in approximately 15 weeks. This model based on the first two entries of a blog contained fewer predictors than in the model for weight loss based on sampling from the entire 15 week period.

Each of the measurement methods used (i.e. word counts, blogging rates, and social support measures) contributed to an understanding of the self-change process within online blog communities. Overall, the level of investment within a blog community was a strong predictor of the degree to which people kept up with the blog and lost weight. For example, more personal blogs (as indicated by the use of first person singular pronouns) were more likely to be associated with weight loss than keeping an online food diary (as indicated by ingestion word use). Given the abundance of empirical literature documenting the importance of keeping frequent food diaries for the weight loss process, this finding was surprising.

There are several reasons why some of the current findings are inconsistent with experimental studies where eating behavior is observed in a controlled lab setting. The time span between variables that have been experimentally manipulated in lab studies (e.g. mood, self-focus, etc.) and their effects (e.g. overeating) were unlikely to have been captured by the large time period sampled in the current study (i.e. 15 weeks). One ecological momentary assessment study of binge eating found increases in negative mood and self-criticism immediately before and immediately after binges. Given the “potentiating” nature of binges (Steiger, Gauvin, Engelberg, Kin, Israel, Wonderlich et al., 2005), it was expected that increases in weight might be detected when negative emotion words and self-focus appear together at high rates during the sampling period.

However, perhaps binges are not associated with vast changes in weight, perhaps the bloggers had taken a few days off the blog to binge, or perhaps the level at which we assessed self-change did not address smaller fluctuations in self-regulation. Future research might assess this by examining language and weight change on an entry by entry level.

Another finding that contradicted expectations was that an increase in cognitive mechanism word use was actually associated with dropout and weight gains rather than with weight loss. Taken together, these findings do not support the limited resource theory of self-regulation. The limited resource theory of self-regulation posits that heightened negative attention to one's self leads to narrowing attention to the immediate environment. In the case of dieters, negative attention to one's self would lead to narrowing attention to food, breaking one's diet, and overeating (Heatherton & Baumeister, 1991). Furthermore, cognitive mechanism words in narratives of past negative events have been shown to increase working memory capacity (e.g., Klein & Boals, 2001). Accordingly, the limited resource theory of self-regulation would hypothesize that the use of cognitive mechanism words would increase the likelihood of staying on a diet (i.e. free up resources to maintain dietary restraint). However, in the dissertation, the use of first person singular pronouns was positively associated with weight loss, while the use of cognitive mechanism words was associated with weight gains, and even dropout from the diet blogs. The relationships between self-awareness, working memory capacity, and instances of self-regulation failure cannot be determined from the analyses here; the current dissertation cannot discount the limited resource

theory of self-regulation since diet bloggers may have been bingeing in private. Future analyses might examine these constructs on an entry by entry level, with special attention to the latency to the next entry when first person singular pronouns and negative emotions words appear at high rates within an entry (suggesting negative self-focus).

Consider also that the number of comments made to and received from other blogs were also strongly associated with weight loss. It would be easier to imagine more personal style blogs leading to more comments received than food diary style blogs, and connections within the online community easier to develop if personal information was being shared. The positive relationship of weight loss and social integration within an online blog community was also observed in the frequency with which bloggers posted entries. Specifically, a greater number of active weeks on the blog (i.e. number of weeks in which there was at least one blog entry), was consistently associated with weight loss. Perhaps the social support received from the blog community acted as a buffer or “resource” even when resources for a given dieter had been depleted by ego threats or other triggers for self-regulation failure. An alternate hypothesis is that those who have a positive outlook and are more open to change to begin with are more likely to comment to others, and to broaden and build their social network (Fredrikson, 2005). Future research in self-change might seek to assess personality traits (e.g. extraversion, openness, self-efficacy) and dispositional mood states (e.g., optimism, excitement) that are associated with successful self-change attempts.

Previous work on disclosure would predict that the best correlate of successful weight loss would be the language people use. All of Pennebaker’s expressive writing

research points to this. The big difference between Pennebaker's research, however, is that writing is done in a solitary context and not publicly shared. Blogging is a very different enterprise. Not only do people share their entries with others but they clearly seek feedback from others (Nardi, Schiano, & Gumbrecht, 2004). Blogging, then, is a social process – perhaps more than it is a self-reflective, and insight-seeking or insight provoking self-change strategy. Over virtually all of these analyses in this study, the magnitude of the goal, and the social engagement measures, including number and types of comments from others as well as the degree of making comments to others, were far more powerful in predicting subsequent weight loss than any of the language markers. Indeed, the relative power of the social variables over the language variables is the most unexpected and most important finding of the dissertation.

IMPLICATIONS

Much of the emphasis in current behavioral and clinical weight loss programs is on educating clients about portion control, food pyramids, nutrition, and the benefits of exercise. Components typically include the monitoring of food intake through the use of food diaries, which have consistently been found to be a significant predictor of weight loss (e.g. Hollis et al., 2008). However, the use of food diaries as measured by the use of ingestion words was not a significant predictor of weight loss on the blogs. In fact, focusing more of the first couple entries on food intake was found to be higher in dropouts than in dedicated bloggers. It cannot be ruled out that bloggers who experienced a great deal of weight loss were keeping food diaries. However, it is important to note that bloggers who lost weight were less likely to report solely on their food intake in their

blog entries. Indeed, successful dieters were more likely to post about personal issues and emotional events in their blog entries. For weight loss blogs in a community of blogs, keeping a food diary may not have been conducive to receiving many noteworthy comments from other bloggers. Indeed, it was the numbers and kinds of comments received that were large determinants of weight loss in the blogs.

Taylor (2007) identified 3 broad types of social support. Informational support is the sharing of information and advice; instrumental support is the sharing of tangible assistance or goods; emotional support is the sharing of warmth and nurturance. No direct evidence for the provision of informational or instrumental support was measured through language variables. However, upon reading the blogs and comments regarding dieting tips and advice, informational support was clearly a large part of the social support offered within the community. In addition, emotional support was a strong influence on the weight loss process; the use of negative emotion words in the comments was predictive of dropping out. So, perhaps those considering a diet blog should consider that the aim is to write the first few entries on topics that will elicit positivity from the blog's audience, and to wait or to refrain on disclosing topics that might shun away commenters or elicit negative comments.

The findings point to the types of social support that might help dieters in behavioral programs or in self-change efforts. Note that there were no formal clinicians, dieticians, or nutritionists on the www.dietdiaries.com (2007) website, but that it was made up of members who were interested in losing weight. The degree to which bloggers were socially integrated was a large determinant of weight loss. Accordingly, behavioral

weight loss programs with online components might provide dieters with more opportunities to comment or discuss personal issues with other dieters through the ability to browse other participants' profiles, make comments, and to receive comments from other members. Indeed, other studies have shown increased commitment to medical treatments and increases in self-efficacy when social support is available (Resnick, Orwig, Magaziner, & Wynne, 2002). Note also that the success of Weight Watchers, a commercial weight loss program that features a venue for social support in the weight loss process, has shown variable but consistently positive results for weight loss (for a review, see Tsai & Wadden, 2005; Womble, Wang, & Wadden, 2002).

Given that making comments to many other bloggers was found to be a predictor of staying on the blog and losing weight in the current study, participants of online diet sites should make an effort to spend some portion of their time on a site commenting to other dieters. It is unknown whether this has the effect of increasing the number of comments received on their own progress, thereby increasing emotional support for one's self, or increasing informational support by reading diet tips, successes, and failures from other dieters. It is likely that a combination of all of these mechanisms led to a perception of social support that facilitated weight loss. Future research can build more complex statistical models of pathways to successful weight loss in order to disentangle these effects. The results of the dissertation point towards the importance of predicting 1) what features of entries predict receiving positive, negative, and many comments, 2) what features predict commenting to others, and 3) which commenters in the social community influence weight loss by other bloggers. Clearly, many more social network analyses

must be conducted in order to uncover the dynamics of what lures people to certain blogs, and what kinds of comments influence weight loss and blogging behavior. In any case, it is encouraging that some of the work involved in losing weight can happen through the development of a supportive online social network, instead of only through time spent in strenuous exercise along with restraint from deliciously fatty foods.

As for what the dissertation reveals about the potential of blogging about weight loss or about starting any self-change endeavor, the results indicate that the mindset upon starting a self-change blog or support program is important. Bloggers who increased in their use of cognitive mechanism words and those who used a high proportion in their first two entries were more likely to gain weight or to dropout early on. Being explicit about perceived (or anticipated) social support instead of using the blog to start contemplating how one should and might change was an effective strategy. Future research might focus on assessments of stages of self-change through the use of computerized text analysis upon intake into behavioral treatment programs to identify those who are ready to benefit, and those who might need more convincing before committing.

LIMITATIONS

The primary limitation of the dissertation concerns the causal pathway: Were bloggers losing weight because they were commenting to other blogs? Or were they commenting to other blogs because they were losing weight? Were extraverted bloggers more likely to comment to others and lose weight? Or does losing weight encourage gregariousness and rampant commenting to other blogs? Future studies might look at

blogs on a more fine-grained level, entry by entry, or create path models for moderating effects. Overall, the relationships between the target variables need to be assessed for what predicts weight loss. The results from the current study suggest that examining what predicts making and receiving comments would be a start. In addition, the results from the cognitive mechanism words predicting dropout and weight gain suggest that future studies might take a closer look at the language associated with the six stages of self-change (Prochaska et al., 1992; 1995). For example, judges could code each blog for the stage of self-change that bloggers were in upon starting and throughout the blog. Language associated with each of the stages could be identified. Then these language variables could be entered into regression analyses to determine which stages of self-change are most associated with dropping out or weight loss. Previous research suggests that those in the precontemplation or contemplation stage of self-change would fare worse in the weight loss process, and the use of cognitive mechanism words in the current dissertation may have reflected thinking in these stages.

Most of the results from the current study are only generalizable to females. However, the Internet has increasingly been used by individuals to find others with their stigmatized illness (Baker, Wagner, Singer, & Bundorf, 2003; Berger, Wagner, and Baker 2005; Davison, Pennebaker, & Dickenson, 2000; Fox and Fallows 2003). Accordingly, males may be likely to use blogs more than face to face support groups to talk about body and weight issues, which have traditionally been associated with female concerns. In the current sample, 3.1% of the blogs were by males. However, for a large portion of the sample, sex could not be determined. It would be interesting to use text

analysis to classify the sex of these indeterminable bloggers based on their word use (e.g., Koppel, Argamon, & Shimoni, 2003), and then to see if this group was different from highly identifiable bloggers in some way. For example, it might be hypothesized that publicly declaring one's identity would lead to accountability or follow-through on goals and dedication to the blog (Cialdini, 1993; Kiesler, Pallak, & Kanouse, 1968).

Another limitation concerns the self-reported weights. As mentioned, previous studies have found high correlations between self-reported recall of weights and weights at the lab, although some variability has been found based on age, gender, and weight (Kuczmarski, Kuczmarski, & Najjar, 1995 ; Perry, Byers, Mokdad, Serdula, & Williamson, 1995). It is unlikely however that bloggers would want to lie about their weight on a given day, considering the associated written entry would often comment on the weight for the day, and considering that the blog was intended as a means to track weight. In fact, many people were eager to report their weight up to two decimal places, and even blogged frequently about the anticipation of going to the doctor's office or clinic just to get on a more precise scale.

The problem with the self-reported weights for future studies that might examine the relationships of language and weight loss at the entry by entry level is that the weights were not reported consistently for some bloggers. Based on our analyses, it may be possible to determine, based on word use, blogging rates, and comments, whether or not the missing weight days were "bad" (weight gain) or "good" (weight loss) days. Indeed, this latter point raises the issue of not having verified any of the information in the corpus

by asking the bloggers directly. However, this non-reactivity is also an advantage to the current study. The tradeoff is in the naturalistic data versus the experimental data.

The final limitation concerns the generalizability of the findings. In this dissertation, only a single diet blog community was examined. Commenting was open only to its members who kept a blog on the community. Although there are other diet blog communities, the one examined in this dissertation is one of the largest sites devoted to dieting. For the purposes of examining self-change, blogs were ideal, since these enabled individuals to be more easily tracked over potentially long periods of time. Online bulletin board message forums, which are common on other diet sites, often contain sporadic questions or comments about common dieting practices, with less information on the self-change process recorded throughout an individual's diet. However, other diet blog communities and individual blogs about weight loss not part of a diet blog community must be examined before generalizations from the current dissertation can be made.

FUTURE DIRECTIONS

The results of the dissertation point to the social and psychological importance of sending and receiving comments within the blog world. Also, since the degree of weight loss by friends has been shown to be significantly associated with that of a participant (e.g. Wing & Jeffery, 1999), future research might assess the relationship of weight loss between commenters and a blogger. Indeed, many more rich analyses can potentially be conducted in order to assess the development and maintenance of the various social networks within the community.

Alternate or additional analyses could be conducted using the blog profile information, or ratings of the pictures posted in the blogs. Also, an analysis for seasons of self-change or days of the week of failure, for example, could be conducted using the date stamps. It has been observed that the time when people most often promise themselves to change is on or around the New Year (e.g. Marlatt & Kaplan, 1972), and that restraint failures often occur on weekends and holidays (Gorin et al., 2004). The blogs can be analyzed for seasonal changes in weight, weight loss, weight gain, dropouts, and new starts. Then, the blogs could be assessed for time periods in which the most weight can be lost (e.g. starting at the beginning of bikini season, and only posting on weekends, etc.).

Another temporal analysis can come from assessing how realistic bloggers set their weight loss goals by going beyond goal discrepancy from starting weight studied in the dissertation, and by looking at the anticipated time frame set for the goals. For example, do bloggers that set goals to be attained within a smaller time frame achieve greater long term weight loss than do bloggers who set goals within a more distant time frame? Construal level theory posits that temporally distant goals generate higher-level or more abstract construals of a goal, while temporally near goals generate lower-level or more concrete features of a goal (Trope & Liberman, 2003). Accordingly, assessing the time frame of weight loss goals, along with identifying times at which goal adjustments are made during blogging may contribute to a better understanding of self-change. Indeed, research suggests that time pressure and time horizons can affect action

implementation, along with persistence and disengagement from goals (see Henderson, Gollwitzer, & Oettingen, 2007).

Experimental studies have shown that reliable and accurate personality impressions can be gleaned from just a thin slice of online information. For example, previous research has shown that simply browsing an individual's homepage can provide reliable and accurate impressions about their personality (Vazire & Gozling, 2004). Other researchers have observed highly reliable and accurate ratings of personality such as openness, narcissism (Back, Schmukle, & Egloff, 2008), and self-esteem (Chang-Schneider & Swann, 2008) from e-mail addresses alone. The latter study found differences in how people were willing to interact with someone based on the e-mail address information alone. Researchers have also found gender information to be commonly embedded in screen names for chat forums (Subrahmanyam, Smahel, & Greenfield, 2006). Since users choose screen names to be reflective of self-perceived traits, perhaps screen names are also chosen in a way to convey motivations to be online, and influence the degree to which users on an interactive online forum interact. Given that most interactions within online forums are initiated by browsing from a list of screen names, choosing a particular screen name to use or to interact with are significant if the goal of a diet blogger is to elicit a particular online experience. In order to assess if weight loss can be predicted based on screen name information alone, the screen names of users can be rated for the degree to which judges' think that they will lose weight or would comment to the screen name user's blog. These ratings can be assessed for their validity (e.g. whether or not judges' ratings are accurate about the starting weight of a

blogger, how much weight they think that a blogger will lose, and how many comments or followers they think the blogger will receive).

Outside of the weight loss domain, the self-change process can be examined in other blog communities devoted to self-change, such as blogging away credit card debt (see Leland, 2007), or positive goals such as training for a marathon. In short, there are many public blogs available with words and other quantitative measures of self-change currently archived on the Internet. This dissertation offers one way to assess the information in blogs to understand and to predict the self-change process aided by computerized text analysis.

With the vast amounts of texts archived on the Internet and with recent advancements in computerized text analysis, the ability to efficiently assess natural language has grown. Previous research has found reliable linguistic markers of psychological states using computerized text analysis tools. Accordingly, these tools have been applied to special interest groups that have come together online, offering a non-obtrusive way to examine the psychological characteristics of people with a certain illness, interest, background, goal, or attitude. Text analyses of online language can tell us how people are relating to their topic, to their audience, and to themselves in ways that go beyond the snapshot offered by brief behavioral assessments in the lab or by questionnaires.

The analyses conducted in the dissertation focused primarily on word count tools. However, many more text analyses can help to better understand dieting. For example, the current study found that the level of social integration was a very strong predictor of

weight loss, and that the number and types of comments received were associated with the degree of weight loss. Future studies might assess the degree to which a blogger and its commenters match in their language style. Language style matching (LSM) is a measure of the degree to which two or more people use the same rate of function words, and has been shown to be a robust marker of engagement and liking (Ireland & Pennebaker, 2009; Niederhoffer & Pennebaker, 2002). Accordingly, LSM might be linked to the degree to which bloggers will stay committed to the blog community. Furthermore, given that weight loss has been found to be similar within a group of diet buddies (Wing & Jeffery, 1999), future research could also test whether LSM is higher between bloggers who are losing weight, and lower between bloggers who break their diets and do not lose as much weight as their social network.

Another promising text analysis method that could contribute to the richness of the findings in the dissertation would be to determine the topics amongst dropouts, successful dieters, and unsuccessful dieters. Instead of developing an elaborate coding scheme for the open-ended nature of blogs, an inductive method could be used to automatically extract themes from the blogs. One such method is the Meaning Extraction Method (MEM; Chung & Pennebaker, 2008; Pennebaker & Chung, 2008). The MEM finds clusters of words that tend to co-occur in a corpus. These clusters tend to form coherent themes that have been shown to produce valid dimensions for a variety of corpora. For example, Chung and Pennebaker (2008) showed that their MEM-derived personality factors were correlated with the BFI in some expected ways (e.g. neurotic people talked more about negative emotions; introverts talked about socializing or not;

extroverts talked about ways of relating to others; conscientious people talked more about school/work). Pennebaker and Chung (2008) showed that their MEM-derived word factors of al-Qaeda statements and interviews across time differentially peaked during the times when those factors/topics were most salient to al-Qaeda's missions. Finally, in an unpublished data set, Chung, Rentfrow and Pennebaker (2009) have found that MEM-derived word factor scores from "This I Believe" essays (i.e. open-ended essays written to address one's personal beliefs) were strongly associated with Census Bureau data and state-level personality traits in expected ways. In sum, MEM-derived factors have been shown to have content validity across multiple domains. For the diet blogs, the MEM could be applied to assess the common topics of a corpus that were used by successful and unsuccessful dieters, and to assess different topics that arose throughout the self-change process. Since the MEM does not require a predefined dictionary, as word count strategies rely on, the MEM can be applied to blogs or self-change texts in any language (e.g. Ramirez-Esparza et al., 2008; Ramirez-Esparza, Chung, Sierra-Otero, & Pennebaker, 2009; Wolf et al., in press).

CONCLUSIONS

Research in psychology has been criticized for avoiding the assessment of behaviors in everyday life. Participants are often brought into the lab under highly controlled conditions. Sometimes they are sent back home, and are asked to fill out a questionnaire on their experiences since their time in the lab. While blogs do not provide a set of standard instructions or highly controlled conditions with which to study a set of behaviors, they provide a window into the everyday lives of regular people (or at least of

online users). In the dissertation, a computerized text analysis of diet blogs in an online community provided a look at the processes underlying self-change in everyday life, and the context in which they unfolded. Some of the processes found to be strongly associated in retrospective accounts of self-change, in experimental lab studies of overeating, and in clinical weight loss treatments were not strong predictors of weight loss in blogs. Given the richness of the findings in the current dissertation, researchers, clinicians, and garden variety dieters are encouraged to go online to learn what affects self-change in everyday life. The dissertation showcased the use of computerized text analysis for understanding the psychology of online users in everyday life.

Appendix A: Instructions to Judges for Ratings of Entries

Read a single entry, and in the same row, answer the items in the headings for each entry. For those items with the "can't tell" option, mark the cell with an "x" if you've tried but really can't tell. For those items that don't have "can't tell" as an option, force yourself to choose.

1. To what degree is this blogger self-focused? (i.e. their attention is directed on themselves or inward)

0 = not at all
1 = very little
2 = somewhat
3 = a lot

Note that for emotions, the question asks whether or not the blogger is expressing that emotion (NOT the degree to which the entry is mostly negative or mostly positive).

2. To what degree is this blogger expressing negative emotions?

0 = not at all
1 = very little
2 = somewhat
3 = a lot

3. To what degree is this blogger inhibiting or suppressing their negative emotions?

x = can't tell
0 = not at all
1 = very little
2 = somewhat
3 = a lot

4. To what degree is this blogger expressing positive emotions?

0 = not at all
1 = very little
2 = somewhat
3 = a lot

5. To what degree is this blogger inhibiting or suppressing their positive emotions?

x = can't tell
0 = not at all

1 = very little
2 = somewhat
3 = a lot

6. To what degree is this blogger making realizations?
(i.e. learning/growing/discovering/gaining insight)

x = can't tell
0 = not at all
1 = very little
2 = somewhat
3 = a lot

7. To what degree is this entry a food intake diary (i.e. a record of the kinds and amounts of food or calories consumed within any given time period)?

0 = not at all
1 = somewhat (i.e. a minority of the entry is a food diary)
2 = mostly (i.e. the majority of the entry is a food diary)
3 = completely (i.e. just a food diary)

8. To what degree does this blogger feel social support from others? (can be from people online or offline)

x = can't tell
0 = not at all
1 = very little
2 = somewhat
3 = a lot

**Appendix B: Correlations of Judges' Ratings of Food Diaries with
LIWC's Biological Mechanisms Category and Subcategories**

LIWC Categories	Food Diary	
	<i>r</i>(103)	<i>p</i>
Biological Processes	.55	.000
Body	.03	.753
Health	-.01	.913
Sexual	-.02	.868
Ingestion	.65	.000

**Appendix C: Correlations of Judges' Ratings of Social Support with
LIWC's Social Category and Subcategories**

LIWC Categories	Social Support	
	<i>r</i>(103)	<i>p</i>
Social	.35	.000
Family	.05	.640
Friends	.27	.006
Humans	.07	.502
Personal Pronouns	.15	.133
I	-.04	.724
We	.17	.090
You	.35	.000
SheHe	.09	.369
They	-.16	.120

**Appendix D: Correlations of Judges' Ratings of Making Realizations
with LIWC's Cognitive Mechanism Subcategories**

LIWC Categories	Making Realizations	
	<i>r</i>(103)	<i>p</i>
Cognitive Mechanisms	.18	.072
Insight	.38	.000
Causal	.24	.013
Discrepancy	-.13	.184
Tentative	.11	.262
Certainty	.18	.077
Inhibition	-.05	.600
Inclusive	.08	.422
Exclusive	-.09	.386

Appendix E: Correlations of Judges' Ratings of Expressing and Suppressing Negative and Positive Emotions with LIWC's Affect Categories and Subcategories

LIWC Categories	Expressing Negative Emotions		Suppressing Negative Emotions		Expressing Positive Emotions		Suppressing Positive Emotions	
	<i>r</i> (103)	<i>p</i>	<i>r</i> (103)	<i>p</i>	<i>r</i> (103)	<i>p</i>	<i>r</i> (103)	<i>p</i>
Affect	-.14	.165	.08	.442	.40	.000	-.08	.435
Negative Emotions	.24	.014	.18	.068	-.15	.121	-.02	.877
Anxiety	.32	.001	.12	.216	-.18	.073	-.03	.768
Anger	.21	.033	.07	.488	-.14	.149	.08	.414
Sadness	-.12	.247	.04	.708	.09	.365	.03	.773
Positive Emotions	-.25	.012	-.01	.958	.48	.000	-.09	.368

Appendix F: Correlations of Judges' Ratings of Self-Focus with Pronoun Use

LIWC Categories	Self-Focus	
	<i>r</i>(103)	<i>p</i>
Personal Pronouns	.05	.624
First Person Singular	.17	.079
I	.07	.491
Me	-.06	.545
My (includes mine, myself)	.23	.021
First Person Plural	-.02	.844

Appendix G: Correlations between Judges' Ratings of Items

Item	Correlations with Coding Categories $r(115)$							
	Self-Focus	Express Negemo	Suppress Negemo	Express Posemo	Suppress Posemo	Making Realizations	Food Diary	Social Support
Self-focus	1.00	.05	-.08	-.23*	.14	-.15	.08	-.22*
Expressing negative emotions		1.00	.44***	-.60***	.11	-.32***	.04	-.44***
Suppressing negative emotions			1.00	-.19*	.12	-.16	.08	-.26**
Expressing positive emotions				1.00	-.06	.57***	-.20	.68***
Suppressing positive emotions					1.00	-.02	.07	-.16+
Making realizations						1.00	.02	.46***
Food diary							1.00	-.22*
Social Support								1.00

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, two-tailed. Empty cells are redundant.

Appendix H: Correlations of Judges' Ratings with all LIWC Categories

LIWC Categories	Examples	Correlations with LIWC Categories $r(103)$							
		Self-focus	Expr. Neg Emo	Suppr. Neg Emo	Expr. Pos Emo	Suppr. Pos Emo	Making Realizations	Food Diary	Social Support
Linguistic Processes									
Words/sentence		-.11	.03	-.07	-.05	-.11	-.17	-.00	-.17
Dictionary words		-.03	.04	-.04	.15	-.03	.22*	-.30**	.01
Words>6 letters		-.03	.04	.01	.00	.03	.15	.15	.09
Total function words		.15	.12	.02	-.02	-.02	.07	-.34***	-.04
Total pronouns	I, them, itself	-.01	.08	-.12	.10	.04	.24	-.35***	.08
Personal pronouns	I, them, her	.05	.03	-.11	.13	-.05	.17	-.36***	.15
1st pers singular	I, me, mine	.17	.05	-.09	-.03	.10	.15	-.20*	-.04
1st pers plural	We, us, our	-.02	.00	.05	.15	-.01	.09	-.12	.17
2nd person	You, your, thou	-.18	-.34***	-.14	.37***	-.19	.14	-.20*	.35***
3rd pers singular	She, her, him	-.12	.18	.05	.03	-.17	-.04	-.09	.09
3rd pers plural	They, their, they'd	.06	.18	-.01	-.15	-.13	-.13	-.15	-.16
Impersonal pronouns	It, it's, those	-.10	.09	-.04	-.02	.14	.17	-.08	-.08
Articles	A, an, the	.08	-.11	.07	-.04	.02	-.03	.09	-.15
Common verbs	Walk, went, see	-.04	.02	-.04	-.15	.03	-.22*	-.20*	-.04
Auxiliary verbs	Am, will, have	.05	.16	.06	-.15	-.03	-.09	-.24*	-.12
Past tense	Went, ran, had	.09	.14	.11	-.11	.01	-.27**	.08	-.04
Present tense	Is, does, hear	-.09	-.10	-.20*	.04	-.06	.12	-.39***	.01
Future tense	Will, gonna	-.03	-.01	.02	-.10	.09	-.07	-.04	-.05
Adverbs	Very, really, quickly	.12	.03	-.02	.06	.07	.03	-.23*	-.02
Prepositions	To, with, above	.15	-.01	.04	.01	-.18	-.04	-.09	-.04
Conjunctions	And, but, whereas	.09	.11	-.05	-.11	.12	.08	-.10	-.04
Negations	No, not, never	.15	.15	-.12	-.15	-.02	-.03	.02	-.20*
Quantifiers	Few, many, much	-.16	-.02	.06	-.03	-.06	-.06	-.02	.02
Numbers	Second, thousand	.05	-.04	.23*	.09	.02	.05	.08	.10
Swear words	Damn, piss, fuck	-.02	.08	-.08	-.03	.24*	-.07	-.05	-.06
Psychological Processes									
Social processes	Mate, talk, they, child	-.28**	-.02	-.04	.23*	-.19*	.03	-.31**	.35**
Family	Daughter, husband, aunt	-.10	.20*	.09	.01	-.07	-.08	-.13	.05
Friends	Buddy, friend, neighbor	-.07	-.02	.12	.16	-.16	.02	.01	.27**
Humans	Adult, baby, boy	-.18	-.01	.01	.12	-.06	.02	-.06	.07
Affective processes	Happy, cried, abandon	-.25*	.14	.08	.40***	-.08	.30**	-.13	.33***
Positive emotion	Love, nice, sweet	-.30**	-.25*	-.01	.48***	-.09	.24*	-.15	.40***
Negative emotion	Hurt, ugly, nasty	.02	.24*	.18	-.15	-.02	.14	.02	-.12
Anxiety	Worried, fearful, nervous	.09	.32**	.12	-.18	-.03	.06	.02	-.19
Anger	Hate, kill, annoyed	.01	.21*	.07	-.14	-.08	-.11	.07	-.13
Sadness	Crying, grief, sad	.02	.12	.04	.09	.03	.20*	-.02	-.01
Cognitive processes	cause, know, ought	-.12	.14	-.04	-.02	.16	.18	-.11	.02
Insight	think, know, consider	-.14	-.09	-.26**	.13	-.02	.38***	-.17	.07
Causation	because, effect, hence	-.02	-.02	-.12	.01	.28**	.24*	.07	-.03
Discrepancy	should, would, could	.01	.11	-.04	-.14	-.01	-.13	-.36***	-.10
Tentative	maybe, perhaps, guess	-.19	-.00	-.02	-.01	.10	.11	.02	.07
Certainty	always, never	-.18	-.02	-.01	.26**	.00	.18	.07	.23
Inhibition	block, constrain, stop	.05	.15	.09	-.03	.08	-.05	-.10	.00
Inclusive	And, with, include	-.02	.23*	.11	-.05	-.05	.08	.15	.07
Exclusive	But, without, exclude	.07	-.01	-.06	-.07	.13	-.09	-.22*	-.10
Perceptual processes	Observing, heard, feeling	-.13	-.21*	-.10	.19*	.02	.20*	.10	.01
See	View, saw, seen	-.11	-.21*	-.08	.21*	.06	.14	.11	-.05
Hear	Listen, hearing	-.06	.14	-.06	-.01	-.04	-.06	-.19	.05
Feel	Feels, touch	-.09	-.21*	-.08	.16	.06	.23*	-.01	.08
Biological processes	Eat, blood, pain	-.01	-.03	-.15	.02	.04	.08	.55***	-.13
Body	Cheek, hands, spit	.18	.04	-.10	.01	-.07	-.17	.03	-.12
Health	Clinic, flu, pill	-.01	.15	-.13	.07	-.06	-.02	-.01	-.07

Sexual	Horny, love, incest	-.05	-.20*	-.11	.28**	-.09	.11	-.02	.32***
Ingestion	Dish, eat, pizza	-.01	-.07	-.09	-.03	.09	.12	.65***	-.11
Relativity	Area, bend, exit, stop	.11	.03	.28**	-.04	.08	-.08	-.10	.01
Motion	Arrive, car, go	.02	.06	.20*	-.05	.12	-.10	-.04	-.11
Space	Down, in, thin	.11	-.01	.19	.03	-.03	-.09	-.09	.05
Time	End, until, season	.01	.00	.13	-.05	.06	.03	-.14	.07
Personal Concerns									
Work	Job, majors, xerox	-.11	-.14	-.17	.01	-.21*	-.08	-.22*	-.01
Achievement	Earn, hero, win	-.05	-.20*	-.04	.24*	-.03	.23*	-.04	.08
Leisure	Cook, chat, movie	-.17	-.11	.02	.12	-.02	.04	-.01	.02
Home	Apartment, kitchen, family	.11	.13	-.03	-.12	-.16	-.28**	-.09	-.05
Money	Audit, cash, owe	-.13	-.01	.02	-.06	.04	-.05	.05	-.15
Religion	Altar, church, mosque	.22*	-.11	-.05	.15	-.05	.08	.05	.18
Death	Bury, coffin, kill	.03	.05	.08	-.13	-.10	-.07	.01	-.13
Spoken categories									
Assent	Agree, OK, yes	-.06	-.03	.05	.22*	-.03	.06	.03	.08
Nonfluencies	Er, hm, umm	-.05	.12	.00	.06	.08	.04	.01	.00
Fillers	Blah, lmean, youknow	-.16	-.11	-.24*	.10	-.04	.22*	-.07	.16

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, two-tailed

Appendix I: Correlations of Percent Body Weight Loss over the 15 Weeks with all LIWC Categories, other Blogging Rates, and Comment Counts

		<i>M</i>	<i>SD</i>	<i>r</i> with % Weight Loss	<i>p</i>	<i>df</i>	
Blog Entries	LIWC CATEGORIES						
	Linguistic Processes						
	Words/sentence	22.75	26.82	.02	.669	186	
	Dictionary words	87.97	4.50	.11	.136	186	
	Words>6 letters	11.94	1.76	-.15	.044	186	
	Numerals	2.17	1.67	-.06	.410	186	
	Total function words	57.56	4.23	.13	.077	186	
	Total pronouns	16.80	2.33	.14	.053	186	
	Personal pronouns	11.67	1.75	.14	.060	186	
	1st pers singular	9.25	1.61	.18	.013	186	
	1st pers plural	.53	.37	-.04	.602	186	
	2nd person	.65	.45	.02	.808	186	
	3rd pers singular	.84	.62	-.06	.403	186	
	3rd pers plural	.41	.21	-.02	.748	186	
	Impersonal pronouns	5.12	.87	.10	.166	186	
	Articles	5.12	.73	.04	.553	186	
	Common verbs	16.42	1.93	.10	.169	186	
	Auxiliary verbs	10.22	1.33	.08	.277	186	
	Past tense	4.44	.84	-.05	.497	186	
	Present tense	9.71	1.66	.08	.298	186	
	Future tense	1.17	.32	.03	.663	186	
	Adverbs	6.07	1.05	.07	.364	186	
	Prepositions	12.12	1.09	-.05	.516	186	
	Conjunctions	6.83	.97	-.04	.625	186	
	Negations	1.87	.41	-.11	.137	186	
	Quantifiers	3.09	.51	.02	.811	186	
	Numbers	.69	.29	.01	.951	186	
	Swear words	.09	.08	-.11	.139	186	
	Psychological Processes				.00		
	Social processes	5.34	2.05	-.01	.929	186	
	Family	.37	.27	.05	.520	186	
	Friends	.15	.11	-.12	.098	186	
	Humans	.43	.25	.03	.642	186	
	Affective processes	5.57	1.27	.20	.007	186	
	Positive emotion	3.99	1.18	.17	.020	186	
	Negative emotion	1.60	.38	.11	.147	186	
	Anxiety	.27	.12	-.03	.649	186	
	Anger	.29	.15	-.03	.719	186	
	Sadness	.59	.24	.19	.011	186	
	Cognitive processes	15.46	1.51	-.02	.785	186	
	Insight	1.88	.41	.05	.535	186	
	Causation	1.26	.28	-.10	.161	186	
	Discrepancy	1.58	.35	-.05	.545	186	
	Tentative	2.55	.52	-.08	.308	186	

Certainty	1.34	.35	.04	.630	186
Inhibition	.45	.13	-.10	.190	186
Inclusive	4.66	.81	-.01	.931	186
Exclusive	2.93	.61	-.04	.554	186
Perceptual processes	2.31	.45	-.08	.256	186
See	.67	.23	-.11	.155	186
Hear	.32	.15	-.01	.945	186
Feel	1.01	.27	.05	.522	186
Biological processes	4.60	1.38	-.08	.339	186
Body	.77	.36	.02	.791	186
Health	.95	.41	.00	.957	186
Sexual	.20	.16	.07	.319	186
Ingestion	3.06	1.39	-.07	.346	186
Relativity	15.77	1.76	.00	.970	186
Motion	2.24	.47	.05	.493	186
Space	5.47	.74	.03	.655	186
Time	8.06	1.10	-.05	.492	186
Personal Concerns					
Work	1.35	.44	-.07	.344	186
Achievement	1.95	.54	.09	.217	186
Leisure	1.52	.45	-.17	.021	186
Home	.55	.29	-.14	.056	186
Money	.46	.20	-.17	.021	186
Religion	.14	.16	.06	.440	186
Death	.04	.04	-.13	.083	186
Spoken categories					
Assent	.47	.33	.06	.422	186
Nonfluencies	.44	.30	.19	.011	186
Fillers	.25	.10	-.01	.922	186
BLOGGING RATES					
Total entries	49.96	25.97	.12	.11	186
Word count per entry	235.72	127.39	-.06	.45	186
Comments					
LIWC CATEGORIES					
Linguistic Processes					
Words/sentence	14.82	3.42	.04	.577	183
Dictionary words	89.82	2.99	-.18	.014	183
Words>6 letters	11.48	1.29	-.17	.022	183
Numerals	.82	.44	.21	.004	183
Total function words	57.91	2.61	-.15	.038	183
Total pronouns	18.00	1.42	-.12	.102	183
Personal pronouns	12.35	1.22	-.10	.102	183
1st pers singular	5.48	1.06	-.23	.002	183
1st pers plural	.65	.28	.04	.613	183
2nd person	5.22	.87	.13	.081	183
3rd pers singular	.52	.39	.01	.945	183
3rd pers plural	.48	.23	-.04	.642	183
Impersonal pronouns	5.66	.59	-.08	.308	183
Articles	5.50	.50	.11	.140	183
Common verbs	17.39	1.24	.10	.191	183
Auxiliary verbs	10.54	.80	.06	.396	183
Past tense	2.41	.56	-.06	.391	183
Present tense	12.39	1.09	.11	.140	183
Future tense	1.22	.27	-.07	.358	183

Adverbs	5.95	.67	-.12	.096	183
Prepositions	11.56	.84	-.23	.002	183
Conjunctions	6.36	.73	.01	.938	183
Negations	1.65	.32	-.18	.016	183
Quantifiers	2.73	.41	-.30	.000	183
Numbers	.65	.19	-.01	.917	183
Swear words	.07	.07	-.03	.688	183
Psychological Processes					
Social processes	10.38	1.61	.07	.363	183
Family	.38	.28	.07	.339	183
Friends	.14	.10	-.04	.606	183
Humans	.52	.23	.10	.159	183
Affective processes	8.61	1.20	.11	.158	183
Positive emotion	6.94	1.27	.11	.156	183
Negative emotion	1.74	.37	-.05	.530	183
Anxiety	.26	.12	-.02	.746	183
Anger	.27	.14	-.16	.032	183
Sadness	.72	.26	.13	.074	183
Cognitive processes	16.28	1.46	-.14	.060	183
Insight	2.38	.34	-.07	.346	183
Causation	1.36	.32	-.08	.274	183
Discrepancy	1.82	.36	-.12	.099	183
Tentative	2.90	.50	-.15	.041	183
Certainty	1.37	.30	-.24	.001	183
Inhibition	.81	.28	.16	.034	183
Inclusive	4.41	.58	-.01	.912	183
Exclusive	2.84	.51	-.10	.171	183
Perceptual processes	2.73	.59	.11	.149	183
See	.77	.50	.18	.016	183
Hear	.63	.21	-.16	.036	183
Feel	1.05	.27	.11	.124	183
Biological processes	3.71	.93	-.13	.070	183
Body	.62	.27	.02	.773	183
Health	.88	.33	-.05	.516	183
Sexual	.42	.22	-.05	.467	183
Ingestion	2.03	.78	-.13	.088	183
Relativity	13.26	1.29	-.15	.05	183
Motion	1.75	.37	-.02	.830	183
Space	5.42	.65	-.01	.859	183
Time	6.22	.83	-.26	.000	183
Personal Concerns					
Work	1.58	.42	-.04	.613	183
Achievement	2.56	.58	.03	.666	183
Leisure	1.36	.35	-.14	.057	183
Home	.32	.20	-.08	.293	183
Money	.39	.22	-.15	.050	183
Religion	.18	.16	.04	.596	183
Death	.04	.05	-.03	.643	183
Spoken categories					
Assent	.63	.29	.07	.365	183
Nonfluencies	.29	.14	-.04	.581	183
Fillers	.39	.15	-.15	.051	183
COMMENT COUNTS					

Comments per entry	2.83	1.38	.12	.11	186
#Different commenters per entry	2.75	1.31	.13	.08	186
Word count of comment per entry	174.67	111.34	.04	.55	186
Weight of commenters	202.92	18.43	.10	.20	183

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, two-tailed

Appendix J: Correlations between Target Variables in the First 15 Weeks

	Goal Discrep	I	Neg emo	Pos emo	Cog mech	Ingest	Entries/ Wk	Active Weeks	Word Count	Social Words	Com Rec'd	Diff Com	WC Com	Negem o Com	Posem o Com	Com to Others	Others Com To
Start Weight	.91***	.10	.09	.15*	.08	-.16*	-.13	-.02	-.13	.20**	.04	.08	.03	-.10	.13	.01	.05
Goal Discrepancy		.14	.09	.14	.09	-.15	-.11	-.02	-.12	.19*	.04	.08	.04	-.09	.13	.04	.07
I			.43***	.24***	.26***	-.23**	-.26***	-.10	-.27***	-.14	-.20**	-.13	-.21**	.18*	-.09	-.11	-.10
Negemo				.06	.20**	-.12	-.29***	-.05	-.21**	-.04	-.22**	-.15*	-.16*	.29***	-.23**	-.08	-.07
Posemo					.10	-.33***	-.02	.11	-.19**	.34	.09	.15*	.02	.00	.25***	.09	.13
Cogmech						-.32***	-.05	-.00	.15*	.23**	-.02	.04	.02	.15*	-.24***	-.03	-.03
Ingestion words							-.05	-.19*	-.16*	-.55***	-.10	-.18*	-.09	-.13	-.05	-.13	-.13
Mean entries per week								.67***	.80***	.23**	.71***	.62***	.67***	.01	-.08	.54***	.53***
Number of active weeks									.50***	.22*	.45***	.45***	.42***	.13	-.11	.34***	.37***
Total word count										.30***	.65***	.57***	.68***	.01	-.21**	.49***	.44***
Social words in blogs											.33***	.39***	.33***	-.00	.10	.21**	.25***
Total comments in sample												.85***	.96***	-.09	-.03	.75***	.70***
#Different commenters													.79***	-.07	.02	.68***	.71***
Word count of comments														-.08	-.15*	.72***	.67***
Negemo in comments															-.16*	-.06	-.06
Posemo in comments																-.02	-.01
#Comments to other blogs																	.93***

Note. Pearson correlations. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$. Empty cells are redundant.

Appendix K: Original Regression for Research Question 1 with All Cases (i.e. including Outliers)

All of the target variables were entered into a stepwise regression equation. The regression equation was significant, $F(4, 181) = 12.64, p = .000$, and accounted for 21.83% of the variance in percentage body weight loss. The discrepancy between starting weight and goal weight, number of active weeks, first person singular pronoun use in the blogs, and the number of comments made to other bloggers were significant predictors of percentage body weight gain. The regression equation was:

$$\% \text{ Body Weight Loss} = - 6.78 + .04 (\text{goal discrepancy}) + .32 (\text{number of active weeks}) + .59 (\text{first person singular pronouns}) + .01 (\text{number of comments made to other bloggers})$$

Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>
Constant	-6.78	2.89		2.35	.020
Goal discrepancy	.04	.01	.35	5.19	.000
Number of active weeks	.32	.14	.16	2.26	.025
First person singular pronouns	.59	.24	.17	2.50	.014
Number of comments to other bloggers	.01	.00	.15	2.17	.031

Note. $F(4, 181) = 12.64, p = .000, R^2 = 21.83\%$. Negative betas indicate that the more a person was high on a dependent variable, the more weight loss.

However, in this analysis, there was a single case with a standardized residual of percentage body weight loss greater than 3 standard deviations above the mean (i.e. 4.03), and so this case was excluded in a similar analysis reported in the dissertation.

Appendix L: Differences Scores of all LIWC Categories, other Blogging Rates, and Comment Counts over the First and Second Half of Blogging

Blog Entries	LIWC CATEGORIES	First Half (Weeks 1 through 7)			Second Half (Weeks 8 through 15)			Paired <i>t</i>	<i>df</i>	<i>p</i>
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>			
	Linguistic Processes									
	Words/sentence	185	22.10	23.49	184	23.59	30.59	1.48	185	.140
	Dictionary words	185	87.63	5.00	184	88.45	4.75	3.34	184	.001
	Words>6 letters	185	12.02	1.70	184	11.81	2.08	1.41	184	.159
	Numerals	185	2.27	1.92	184	2.02	1.74	2.46	184	.015
	Total function words	185	57.21	4.52	184	58.12	4.64	3.69	184	.000
	Total pronouns	185	16.65	2.40	184	16.97	2.61	1.77	184	.078
	Personal pronouns	185	11.59	1.83	184	11.81	1.95	1.37	184	.173
	1st pers singular	185	9.33	1.67	184	9.17	1.90	1.77	184	.078
	1st pers plural	185	.48	.39	184	.58	.45	3.27	184	.001
	2nd person	185	.64	.46	184	.70	.61	1.70	184	.090
	3rd pers singular	185	.75	.60	184	.94	.81	3.76	184	.000
	3rd pers plural	185	.39	.22	184	.42	.32	1.05	184	.294
	Impersonal pronouns	185	5.06	.88	184	5.16	1.07	1.26	184	.210
	Articles	185	5.13	.82	184	5.08	.89	1.08	184	.283
	Common verbs	185	16.21	2.02	184	16.78	2.24	4.68	184	.000
	Auxiliary verbs	185	10.08	1.39	184	10.51	1.65	4.40	184	.000
	Past tense	185	4.43	.98	184	4.38	1.01	.40	184	.688
	Present tense	185	9.55	1.77	184	10.09	1.91	4.56	184	.000
	Future tense	185	1.15	.38	184	1.15	.49	.16	184	.874
	Adverbs	185	5.99	1.07	184	6.28	1.30	3.94	184	.000
	Prepositions	185	12.10	1.18	184	12.13	1.38	.75	184	.455
	Conjunctions	185	6.83	1.06	184	6.83	1.12	.61	184	.545
	Negations	185	1.84	.44	184	1.88	.52	.49	184	.624
	Quantifiers	185	3.08	.58	184	3.15	.61	1.16	184	.250
	Numbers	185	.71	.32	184	.66	.35	2.31	184	.022
	Swear words	185	.09	.09	184	.08	.08	1.20	184	.232
	Psychological Processes									
	Social processes	185	5.10	2.05	184	5.66	2.34	4.52	184	.000
	Family	185	.36	.29	184	.39	.32	1.34	184	.181
	Friends	185	.14	.12	184	.16	.15	1.69	184	.093
	Humans	185	.41	.25	184	.43	.35	.66	184	.513
	Affective processes	185	5.51	1.29	184	5.69	1.47	2.48	184	.014
	Positive emotion	185	3.94	1.20	184	4.08	1.35	1.47	184	.143
	Negative emotion	185	1.59	.43	184	1.63	.57	1.28	184	.201
	Anxiety	185	.26	.13	184	.30	.24	2.21	184	.028
	Anger	185	.29	.16	184	.28	.18	1.51	184	.132
	Sadness	185	.60	.27	184	.60	.32	.87	184	.386
	Cognitive processes	185	15.35	1.61	184	15.59	1.72	1.55	184	.123
	Insight	185	1.85	.46	184	1.93	.62	1.84	184	.067
	Causation	185	1.24	.32	184	1.30	.41	1.74	184	.083
	Discrepancy	185	1.53	.37	184	1.61	.55	1.32	184	.190
	Tentative	185	2.55	.59	184	2.52	.61	-1.26	184	.210

Certainty	185	1.33	.39	184	1.36	.46	.54	184	.589
Inhibition	185	.46	.17	184	.43	.21	1.77	184	.078
Inclusive	185	4.65	.86	184	4.64	1.03	.30	184	.764
Exclusive	185	2.88	.63	184	2.96	.73	1.23	184	.218
Perceptual processes	185	2.35	.47	184	2.27	.64	1.88	184	.062
See	185	.68	.26	184	.67	.35	.57	184	.567
Hear	185	.31	.16	184	.32	.21	.93	184	.352
Feel	185	1.03	.30	184	1.02	.44	.19	184	.853
Biological processes	185	4.81	1.44	184	4.21	1.55	5.25	184	.000
Body	185	.77	.42	184	.74	.44	1.06	184	.290
Health	185	.96	.47	184	.91	.45	1.24	184	.218
Sexual	185	.19	.15	184	.20	.21	.35	184	.728
Ingestion	185	3.30	1.45	184	2.65	1.47	6.43	184	.000
Relativity	185	15.67	1.96	184	15.92	2.02	2.35	184	.020
Motion	185	2.20	.54	184	2.26	.56	.72	184	.474
Space	185	5.39	.86	184	5.52	.96	2.29	184	.023
Time	185	8.08	1.23	184	8.15	1.33	1.42	184	.159
Personal Concerns									
Work	185	1.33	.50	184	1.42	.60	2.16	184	.032
Achievement	185	2.01	.63	184	1.97	.67	.58	184	.564
Leisure	185	1.53	.49	184	1.49	.59	.95	184	.343
Home	185	.55	.29	184	.55	.35	.36	184	.720
Money	185	.43	.22	184	.50	.37	2.14	184	.033
Religion	185	.13	.17	184	.16	.22	2.20	184	.029
Death	185	.04	.05	184	.06	.23	.94	184	.350
Spoken categories									
Assent	185	.47	.36	184	.46	.37	.49	184	.627
Nonfluencies	185	.43	.31	184	.49	.42	1.85	184	.066
Fillers	185	.25	.11	184	.26	.24	.39	184	.700
BLOGGING RATES									
Total entries	186	28.37	14.14	186	21.59	14.61	7.49	185	.000
Word count per entry	186	233.38				136.7	.51	185	.610
			128.05	186	230.69	0			

Comments	LIWC CATEGORIES									
	Linguistic Processes									
	Words/sentence	182	15.01	4.48	180	14.74	5.98	.62	178	.533
	Dictionary words	182	89.86	3.20	180	89.83	3.63	.16	178	.873
	Words>6 letters	182	11.50	1.44	180	11.35	1.79	1.05	178	.295
	Numerals	182	2.00	1.75	180	.83	.60	.53	178	.597
	Total function words	182	57.91	2.77	180	57.68	3.56	.70	178	.488
	Total pronouns	182	17.93	1.51	180	18.05	2.15	.99	178	.324
	Personal pronouns	182	12.29	1.25	180	12.38	1.69	1.02	178	.309
	1st pers singular	182	5.45	1.14	180	5.46	1.38	.36	178	.718
	1st pers plural	182	.66	.36	180	.61	.40	1.01	178	.314
	2nd person	182	5.21	1.05	180	5.34	1.32	1.24	178	.216
	3rd pers singular	182	.49	.50	180	.53	.55	.60	178	.547
	3rd pers plural	182	.48	.27	180	.44	.32	1.62	178	.107
	Impersonal pronouns	182	5.63	.70	180	5.67	1.06	.40	178	.692
	Articles	182	5.52	.64	180	5.41	1.00	1.26	178	.210
	Common verbs	182	17.27	1.28	180	17.83	2.06	3.51	178	.001
	Auxiliary verbs	182	10.49	.90	180	10.64	1.29	3.51	178	.001
	Past tense	182	2.36	.65	180	2.45	.90	1.03	178	.306
	Present tense	182	12.34	1.15	180	12.81	1.92	3.27	178	.001

Future tense	182	1.23	.33	180	1.18	.43	1.45	178	.150
Adverbs	182	5.91	.75	180	5.99	1.17	.83	178	.407
Prepositions	182	11.61	.99	180	11.51	1.23	1.08	178	.284
Conjunctions	182	6.39	.85	180	6.20	1.07	2.24	178	.026
Negations	182	1.66	.42	180	1.65	.56	.47	178	.636
Quantifiers	182	2.78	.49	180	2.67	.73	1.82	178	.071
Numbers	182	.65	.24	180	.62	.38	.91	178	.363
Swear words	182	.07	.07	180	.08	.16	1.35	178	.179
Psychological Processes									
Social processes	182	10.32	1.86	180	10.46	2.12	.97	178	.334
Family	182	.34	.28	180	.42	.40	2.70	178	.008
Friends	182	.14	.13	180	.13	.15	.33	178	.741
Humans	182	.49	.25	180	.54	.37	1.74	178	.084
Affective processes	182	8.55	1.36	180	9.03	2.09	2.89	178	.004
Positive emotion	182	6.92	1.43	180	7.25	2.14	2.02	178	.045
Negative emotion	182	1.71	.44	180	1.81	.56	1.89	178	.060
Anxiety	182	.24	.14	180	.28	.23	1.81	178	.072
Anger	182	.28	.18	180	.27	.27	.25	178	.799
Sadness	182	.72	.30	180	.74	.41	.43	178	.666
Cognitive processes	182	16.34	1.69	180	16.10	2.16	1.47	178	.144
Insight	182	2.36	.44	180	2.35	.69	.11	178	.915
Causation	182	1.34	.36	180	1.34	.48	.30	178	.764
Discrepancy	182	1.83	.47	180	1.83	.56	.20	178	.839
Tentative	182	2.92	.59	180	2.91	.81	.24	178	.812
Certainty	182	1.35	.39	180	1.37	.50	.34	178	.732
Inhibition	182	.83	.33	180	.81	.46	.61	178	.544
Inclusive	182	4.45	.68	180	4.33	.90	1.45	178	.148
Exclusive	182	2.86	.58	180	2.81	.74	1.05	178	.296
Perceptual processes	182	2.71	.67	180	2.70	1.04	.25	178	.296
See	182	.76	.53	180	.76	.68	.03	178	.980
Hear	182	.62	.24	180	.65	.37	.82	178	.416
Feel	182	1.05	.33	180	1.05	.53	.02	178	.987
Biological processes	182	3.88	.98	180	3.61	1.41	4.18	178	.000
Body	182	.64	.31	180	.60	.44	1.16	178	.246
Health	182	.91	.39	180	.88	.54	.72	178	.472
Sexual	182	.41	.24	180	.48	.48	1.78	178	.076
Ingestion	182	2.21	.89	180	1.69	1.11	5.80	178	.000
Relativity	182	13.25	1.48	180	13.39	2.25	.63	178	.528
Motion	182	1.74	.47	180	1.80	.74	.92	178	.361
Space	182	5.45	.75	180	5.35	1.11	1.20	178	.233
Time	182	6.24	.94	180	6.27	1.44	.19	178	.847
Personal Concerns									
Work	182	1.62	.52	180	1.58	.64	.93	178	.356
Achievement	182	2.62	.70	180	2.57	1.03	.58	178	.566
Leisure	182	1.38	.46	180	1.39	.78	.32	178	.746
Home	182	.30	.22	180	.32	.26	1.51	178	.134
Money	182	.39	.23	180	.40	.36	.75	178	.456
Religion	182	.17	.16	180	.21	.29	1.85	178	.065
Death	182	.04	.05	180	.04	.10	.81	178	.420
Spoken categories									
Assent	182	.61	.30	180	.68	.51	1.84	178	.067
Nonfluencies	182	.28	.16	180	.31	.26	.89	178	.373
Fillers	182	.41	.20	180	.36	.20	2.43	178	.016

COMMENT COUNTS

Comments per entry	186	2.40	1.46	186	2.47	1.87	3.37	185	.001	
#Different commenters per entry	186	3.29	1.89	186	2.90	1.89	3.44	185	.001	
Word count of comments per entry	186	212.78	157.79	186	188.42	169.0	2.40	185	.017	
Weight of commenters	182	200.28	22.28	177	195.63	3	23.76	2.02	176	.045

Tentative	185	-.06	.68	-.02	.757	185
Certainty	185	.02	.42	-.03	.700	185
Inhibition	185	-.03	.26	-.05	.504	185
Inclusive	185	-.02	.96	.04	.631	185
Exclusive	185	.06	.68	-.06	.431	185
Perceptual processes	185	-.08	.61	-.03	.731	185
See	185	-.02	.38	.04	.617	185
Hear	185	.01	.19	.03	.710	185
Feel	185	-.01	.47	-.02	.751	185
Biological processes	185	-.54	1.40	-.11	.148	185
Body	185	-.04	.49	-.15	.037	185
Health	185	-.05	.53	-.07	.316	185
Sexual	185	.00	.18	-.04	.607	185
Ingestion	185	-.59	1.25	-.05	.530	185
Relativity	185	.38	2.18	.04	.640	185
Motion	185	.03	.58	.02	.747	185
Space	185	.16	.96	.02	.750	185
Time	185	.17	1.60	.01	.852	185
Personal Concerns						
Work	185	.12	.73	-.12	.112	185
Achievement	185	.07	1.71	-.04	.567	185
Leisure	185	-.04	.55	.09	.204	185
Home	185	-.01	.28	.05	.503	185
Money	185	.06	.37	-.06	.422	185
Religion	185	.03	.20	-.08	.284	185
Death	185	.02	.24	-.05	.467	185
Spoken categories						
Assent	185	-.01	.28	-.02	.805	185
Nonfluencies	185	.08	.60	-.04	.550	185
Fillers	185	.01	.24	-.03	.687	185
BLOGGING RATES						
Total entries	186	-6.78	12.35	.10	.170	186
Word count per entry	186	-2.69	71.62	.08	.261	186
Comments						
LIWC CATEGORIES						
Linguistic Processes						
Words/sentence	179	-.32	6.77	-.15	.047	179
Dictionary words	179	-.05	4.28	.02	.804	179
Words>6 letters	179	-.16	2.01	.01	.885	179
Numerals	179	.03	.69	-.04	.624	179
Total function words	179	-.21	3.98	.02	.828	179
Total pronouns	179	.17	2.25	-.02	.832	179
Personal pronouns	179	.13	1.71	.09	.214	179
1st pers singular	179	.04	1.54	.04	.600	179
1st pers plural	179	-.04	.53	-.01	.861	179
2nd person	179	.14	1.52	.03	.657	179
3rd pers singular	179	.03	.67	.05	.536	179
3rd pers plural	179	-.04	.35	.07	.354	179
Impersonal pronouns	179	.04	1.21	-.16	.030	179
Articles	179	-.11	1.16	.24	.001	179
Common verbs	179	.56	2.13	.01	.927	179
Auxiliary verbs	179	.17	1.45	.07	.332	179
Past tense	179	.08	1.01	.15	.042	179
Present tense	179	.49	2.01	-.10	.196	179

Future tense	179	-.05	.49	.11	.144	179
Adverbs	179	.07	1.21	-.03	.688	179
Prepositions	179	-.11	1.40	-.01	.915	179
Conjunctions	179	-.20	1.18	-.17	.022	179
Negations	179	-.02	.66	.00	.966	179
Quantifiers	179	-.11	.82	-.01	.879	179
Numbers	179	-.03	.44	.04	.554	179
Swear words	179	.02	.17	-.08	.280	179
Psychological Processes						
Social processes	179	.17	2.35	.06	.428	179
Family	179	.08	.39	.06	.410	179
Friends	179	.00	.19	.03	.716	179
Humans	179	.05	.40	.14	.057	179
Affective processes	179	.46	2.15	.00	.969	179
Positive emotion	179	.33	2.16	-.01	.948	179
Negative emotion	179	.09	.66	.02	.841	179
Anxiety	179	.04	.27	.00	.994	179
Anger	179	-.01	.34	.01	.896	179
Sadness	179	.02	.47	.07	.351	179
Cognitive processes	179	-.25	2.24	-.10	.168	179
Insight	179	-.01	.83	-.04	.587	179
Causation	179	-.01	.51	-.10	.169	179
Discrepancy	179	.01	.67	-.04	.558	179
Tentative	179	-.02	.97	-.07	.377	179
Certainty	179	.02	.65	.02	.765	179
Inhibition	179	-.02	.50	.02	.784	179
Inclusive	179	-.11	1.00	-.10	.190	179
Exclusive	179	-.06	.77	-.06	.417	179
Perceptual processes	179	-.02	1.23	-.13	.076	179
See	179	.00	.86	-.08	.281	179
Hear	179	.03	.43	-.05	.538	179
Feel	179	.00	.62	-.16	.037	179
Biological processes	179	-.46	1.46	-.05	.502	179
Body	179	-.04	.49	-.08	.303	179
Health	179	-.03	.56	-.09	.250	179
Sexual	179	.07	.51	-.05	.506	179
Ingestion	179	-.52	1.20	.04	.601	179
Relativity	179	.11	2.39	.14	.070	179
Motion	179	.05	.79	.15	.041	179
Space	179	-.11	1.28	.07	.377	179
Time	179	.02	1.47	.05	.470	179
Personal Concerns						
Work	179	-.05	.70	-.01	.902	179
Achievement	179	-.05	1.16	.01	.848	179
Leisure	179	.02	.94	-.02	.804	179
Home	179	.03	.29	.03	.658	179
Money	179	.02	.36	-.02	.804	179
Religion	179	.04	.29	-.07	.368	179
Death	179	.01	.11	-.03	.726	179
Spoken categories						
Assent	179	.07	.52	-.02	.785	179
Nonfluencies	179	.02	.27	-.01	.874	179
Fillers	179	-.05	.29	.02	.756	179

COMMENT COUNTS				.00		
Comments per entry	186	-.39	1.59	.08	.279	186
#Different commenters per entry	186	-.39	1.55	.09	.244	186
Word count of comments per entry	186	-.24.36	138.46	.02	.820	186
Weight of commenters	177	-4.21	27.76	.00	.417	177

Note. Negative correlations indicate that the more a person increased on a dependent variable in the second half of blogging, the more weight loss over the 15 week period. Positive correlations indicate that the more a person increased on a dependent variable in the second half of blogging, the more weight gain over the 15 week period.

Appendix N: Correlations between Difference Scores across Halves of Blogging for the Target Variables

	Goal Discrep	I	Neg emo	Pos emo	Cog mech	Ingest	Entries/Wk	Active Weeks	Word Count	Social Words	Com Rec'd	Diff Com	WC Com	Negemo Com	Posemo Com	Com to Others	Others Com To
Start Weight	.91***	-.19**	.03	.03	.05	-.11	-.13	.08	-.05	.13	-.05	.02	-.11	-.04	.13	.01	.06
Goal Discrepancy		.14	.05	.03	.00	-.13	-.11	.06	-.03	.15	-.06	.02	-.12	-.02	.17*	.01	.07
I			.09	-.01	.09	-.02	-.07	-.08	-.03	-.25***	-.10	-.05	-.04	.06	-.02	-.08	-.06
Negemo				-.33***	.06	-.13	-.08	-.05	-.10	-.04	-.09	-.07	-.01	.28***	-.19*	-.03	-.06
Posemo					.01	.00	.03	.10	.05	.05	.04	.12	.00	-.26***	.09	.06	.11
Cogmech						-.14	-.03	.05	.10	.30***	-.01	-.01	.01	.02	.00	-.02	-.05
Ingestion words							.11	.06	.06	-.45	.17	.15*	.13	.03	-.04	.10	.16*
Mean entries per week								.55***	.40***	-.03	.49***	.54***	.44***	-.15*	-.12	.20**	.27***
Number of active weeks									.48***	.08	.38***	.48***	.33***	-.11	-.19**	.18*	.30***
Total word count										.19*	.64***	.55***	.60***	-.08	-.10	.40***	.39***
Social words in blogs											.03	.04	.02	-.09	.16*	.00	-.03
Total comments in sample												.77***	.93***	-.08	-.05	.64***	.59***
#Different commenters													.67***	-.13	-.05	.49***	.57***
Word count of comments														-.03	-.16*	.61***	.54***
Negemo in comments															-.17*	-.08	-.06
Posemo in comments																-.01	.01
#Comments to other blogs																	.86***

Note. Pearson correlations. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$. Empty cells are redundant. Positive correlations indicate that the greater the increase in one variable over the halves of blogging, the greater the increase in the other variable. Negative correlations indicate that the greater the increase in one variable over the halves of blogging, the greater the decrease in the other variable.

Appendix O: Original Regression for Research Question 2 with All Cases (i.e. including Outliers)

The regression equation to predict percent body weight loss over the course of the 15 weeks using difference scores in the variables in addition to starting weight and goal discrepancy at the very beginning of was significant $F(3, 182) = 16.71, p = .000$, and accounted for 21.60% of the variance. The regression equation was:

Percentage body weight loss = .03 + .0004 (goal discrepancy) + .0006 (difference in comments received) - .01 (difference in cognitive mechanism words)

Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>
Constant	.03	.01		4.53	.000
Goal discrepancy	.0004	.00	.38	5.84	.000
Difference in comments received	.0002	.00	.21	3.15	.002
Difference in cognitive mechanism words	-.01	.00	-.19	2.83	.005

Note. $F(3, 182) = 16.71, p = .000, R^2 = .22$. Positive betas indicate that the more a blogger increased in blogging from the first to the second half of blogging on that target variable, the greater the weight loss. Negative betas indicate that the more a blogger decreased in blogging from the first to the second half of blogging, the greater the weight loss. However, for goal discrepancy, the positive beta simply indicates that the larger the discrepancy between starting weight and goal weight, the more weight loss.

However, in this analysis, there was one case with a standardized residual of percentage body weight loss much greater than 3 standard deviations above the mean (i.e. 3.83), and so this case was excluded in a similar analysis reported in the dissertation.

Appendix P: Mean Differences between Dropouts and Dedicated Bloggers on all LIWC Categories for the First Two Entries

Blog Entries	LIWC CATEGORY	Dropouts			Dedicated Bloggers			<i>t</i>	<i>p</i>	<i>df</i>
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>			
	Linguistic Processes									
	Words/sentence	431	22.46	23.53	174	19.12	21.88	1.61	.108	603
	Dictionary words	431	88.45	7.61	174	88.58	5.76	.21	.834	603
	Words>6 letters	431	11.86	3.01	174	12.18	2.51	1.24	.214	603
	Numerals	431	2.43	2.80	174	2.38	2.32	.21	.833	603
	Total function words	431	57.33	7.57	174	57.73	5.59	.64	.526	603
	Total pronouns	431	17.10	3.98	174	17.12	3.25	.06	.953	603
	Personal pronouns	431	12.12	2.93	174	12.06	2.37	.23	.822	603
	1st pers singular	431	10.83	2.72	174	10.36	2.29	2.02	.043	603
	1st pers plural	431	.27	.47	174	.36	.55	1.93	.054	603
	2nd person	431	.46	.60	174	.58	.74	2.22	.027	603
	3rd pers singular	431	.34	.71	174	.51	.87	2.43	.016	603
	3rd pers plural	431	.21	.34	174	.24	.34	1.23	.220	603
	Impersonal pronouns	431	4.98	1.94	174	5.06	1.59	.46	.648	603
	Articles	431	4.70	1.54	174	4.77	1.33	.49	.628	603
	Common verbs	431	16.37	3.20	174	16.38	2.62	.02	.985	603
	Auxiliary verbs	431	10.16	2.47	174	10.17	2.06	.03	.974	603
	Past tense	431	3.75	2.01	174	3.83	1.73	.48	.632	603
	Present tense	431	10.39	3.16	174	10.26	2.62	.49	.627	603
	Future tense	431	1.14	.78	174	1.18	.83	.49	.625	603
	Adverbs	431	5.92	1.90	174	6.06	1.92	.83	.408	603
	Prepositions	431	12.28	2.28	174	12.37	1.79	.47	.638	603
	Conjunctions	431	6.87	1.95	174	6.84	1.48	.20	.840	603
	Negations	431	1.63	.85	174	1.64	.80	.15	.882	603
	Quantifiers	431	2.86	1.22	174	2.99	.97	1.30	.193	603
	Numbers	431	.80	.69	174	.87	.54	1.22	.223	603
	Swear words	431	.09	.23	174	.06	.17	1.53	.128	603
	Psychological Processes									
	Social processes	431	3.99	2.57	174	4.46	2.62	2.05	.041	603
	Family	431	.30	.44	174	.34	.46	.93	.408	603
	Friends	431	.13	.28	174	.15	.23	1.89	.235	603
	Humans	431	.38	.49	174	.38	.42	.06	.951	603
	Affective processes	431	5.73	1.92	174	5.37	1.71	2.15	.032	603
	Positive emotion	431	3.83	1.72	174	3.68	1.65	.95	.342	603
	Negative emotion	431	1.94	1.13	174	1.69	.81	2.66	.008	603
	Anxiety	431	.29	.42	174	.24	.30	1.51	.131	603
	Anger	431	.32	.50	174	.28	.38	.80	.427	603
	Sadness	431	.88	.68	174	.77	.60	1.88	.060	603
	Cognitive processes	431	15.53	3.41	174	15.55	2.63	.05	.961	603
	Insight	431	2.00	1.06	174	1.88	.85	1.32	.187	603
	Causation	431	1.46	.90	174	1.40	.75	.70	.482	603
	Discrepancy	431	1.64	.94	174	1.63	.92	.03	.973	603
	Tentative	431	2.50	1.26	174	2.59	1.00	.83	.410	603
	Certainty	431	1.27	.83	174	1.32	.75	.79	.429	603

Inhibition	431	.49	.53	174	.50	.44	.18	.856	603
Inclusive	431	4.53	1.57	174	4.56	1.43	.23	.819	603
Exclusive	431	2.77	1.28	174	2.87	1.14	.90	.370	603
Perceptual processes	431	2.54	1.22	174	2.50	1.01	.33	.744	603
See	431	.65	.68	174	.65	.56	.14	.889	603
Hear	431	.23	.32	174	.27	.31	1.44	.151	603
Feel	431	1.37	.86	174	1.29	.74	1.01	.314	603
Biological processes	431	5.82	2.92	174	5.43	2.16	1.58	.114	603
Body	431	.68	.71	174	.78	.73	1.45	.147	603
Health	431	1.19	.91	174	1.20	.81	.11	.912	603
Sexual	431	.21	.38	174	.16	.23	1.38	.167	603
Ingestion	431	4.28	2.72	174	3.84	2.01	1.96	.050	603
Relativity	431	15.16	3.55	174	15.84	2.94	2.24	.025	603
Motion	431	1.90	1.04	174	1.98	1.00	.84	.403	603
Space	431	5.20	1.60	174	5.37	1.35	1.21	.228	603
Time	431	8.07	2.55	174	8.49	2.15	1.90	.058	603
Personal Concerns									
Work	431	1.38	1.01	174	1.46	.88	.82	.413	603
Achievement	431	2.82	1.37	174	2.57	1.11	2.14	.033	603
Leisure	431	1.35	.99	174	1.40	.78	.56	.574	603
Home	431	.38	.49	174	.41	.56	.78	.438	603
Money	431	.31	.42	174	.37	.42	1.85	.064	603
Religion	431	.13	.31	174	.09	.21	1.42	.157	603
Death	431	.04	.13	174	.03	.08	1.18	.239	603
Spoken categories									
Assent	431	.34	.47	174	.26	.38	1.79	.075	603
Nonfluencies	431	.41	.47	174	.41	.47	.01	.990	603
Fillers	431	.24	.33	174	.23	.28	.26	.797	603

Comments LIWC CATEGORIES

Linguistic Processes									
Words/sentence	358	15.12	11.85	146	17.18	19.00	1.47	.143	502
Dictionary words	358	91.20	3.01	146	90.44	5.38	2.02	.044	502
Words>6 letters	358	12.29	2.93	146	12.31	2.84	.07	.946	502
Numerals	358	.88	1.02	146	.86	1.00	.21	.833	502
Total function words	358	58.56	3.95	146	57.82	5.18	1.73	.084	502
Total pronouns	358	17.87	3.12	146	17.61	2.93	.86	.388	502
Personal pronouns	358	12.37	2.49	146	12.29	2.49	.34	.737	502
1st pers singular	358	4.95	2.35	146	5.15	2.20	.86	.388	502
1st pers plural	358	.85	.92	146	.74	.75	1.23	.218	502
2nd person	358	5.98	2.17	146	5.70	2.29	1.31	.192	502
3rd pers singular	358	.25	.57	146	.34	.75	1.55	.121	502
3rd pers plural	358	.34	.48	146	.35	.47	.33	.744	502
Impersonal pronouns	358	5.50	1.76	146	5.32	1.76	1.03	.301	502
Articles	358	5.33	1.64	146	5.36	1.46	.23	.821	502
Common verbs	358	17.12	2.65	146	16.92	2.87	.75	.451	502
Auxiliary verbs	358	10.35	2.20	146	10.29	2.27	.25	.803	502
Past tense	358	2.15	1.41	146	2.05	1.19	.76	.448	502
Present tense	358	12.44	2.36	146	12.37	2.43	.28	.779	502
Future tense	358	1.28	.82	146	1.30	.76	.30	.765	502
Adverbs	358	5.76	1.74	146	5.94	1.89	1.02	.309	502
Prepositions	358	12.25	2.01	146	11.88	2.33	1.78	.076	502
Conjunctions	358	6.59	1.77	146	6.51	1.57	.45	.654	502
Negations	358	1.41	1.06	146	1.27	.87	1.45	.148	502

Quantifiers	358	3.08	1.36	146	2.83	1.37	1.93	.055	502
Numbers	358	.68	.66	146	.66	.65	.29	.770	502
Swear words	358	.04	.16	146	.03	.09	1.01	.315	502
Psychological Processes									
Social processes	358	11.65	3.19	146	11.69	3.50	.12	.906	502
Family	358	.22	.46	146	.33	.65	2.04	.042	502
Friends	358	.15	.28	146	.14	.30	.30	.761	502
Humans	358	.40	.49	146	.47	.57	1.51	.132	502
Affective processes	358	9.34	3.06	146	9.24	2.98	.36	.722	502
Positive emotion	358	7.80	3.19	146	7.94	3.15	.46	.649	502
Negative emotion	358	1.61	.97	146	1.39	.86	2.44	.015	502
Anxiety	358	.21	.35	146	.14	.25	2.17	.030	502
Anger	358	.20	.34	146	.19	.30	.39	.700	502
Sadness	358	.83	.76	146	.72	.65	1.54	.125	502
Cognitive processes	358	16.92	3.18	146	16.46	2.96	1.50	.134	502
Insight	358	2.50	1.25	146	2.58	1.21	.62	.536	502
Causation	358	1.39	.87	146	1.32	.85	.77	.440	502
Discrepancy	358	1.84	1.11	146	1.70	1.01	1.40	.163	502
Tentative	358	3.12	1.37	146	2.97	1.43	1.08	.279	502
Certainty	358	1.31	.83	146	1.37	.81	.74	.458	502
Inhibition	358	.75	.67	146	.70	.64	.73	.468	502
Inclusive	358	4.96	1.71	146	4.78	1.57	1.12	.263	502
Exclusive	358	2.64	1.31	146	2.64	1.36	.03	.974	502
Perceptual processes	358	2.66	1.40	146	2.82	1.43	1.11	.268	502
See	358	.82	.77	146	.92	1.06	1.18	.237	502
Hear	358	.50	.53	146	.64	.66	2.55	.011	502
Feel	358	1.10	.87	146	1.03	.72	.90	.369	502
Biological processes	358	3.80	2.09	146	3.44	1.73	1.84	.066	502
Body	358	.41	.55	146	.44	.54	.71	.481	502
Health	358	.86	.87	146	.79	.70	.92	.359	502
Sexual	358	.29	.44	146	.33	.37	.83	.407	502
Ingestion	358	2.46	1.82	146	2.09	1.43	2.19	.029	502
Relativity	358	13.27	3.25	146	12.60	3.01	2.13	.034	502
Motion	358	1.88	1.18	146	1.67	.89	1.93	.055	502
Space	358	5.60	1.73	146	5.43	1.47	1.05	.293	502
Time	358	6.19	2.29	146	6.03	2.23	.69	.490	502
Personal Concerns									
Work	358	1.62	1.11	146	1.74	1.11	1.11	.269	502
Achievement	358	3.15	1.58	146	2.92	1.40	1.54	.125	502
Leisure	358	1.19	.88	146	1.30	.92	1.18	.240	502
Home	358	.21	.39	146	.22	.35	.19	.851	502
Money	358	.28	.43	146	.31	.62	.51	.610	502
Religion	358	.14	.34	146	.18	.55	.93	.351	502
Death	358	.01	.09	146	.03	.09	1.59	.113	502
Spoken categories									
Assent	358	.39	.54	146	.43	.59	.84	.402	502
Nonfluencies	358	.25	.39	146	.29	.48	.86	.388	502
Fillers	358	.33	.39	146	.45	.57	2.58	.010	502
COMMENT COUNTS									
Weight of commenters	375	198.31	43.05	144	201.02	35.27	.67	.50	517

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, two-tailed

Appendix Q: Correlations between Target Variables within the First Two Entries by Dropouts

	Goal Discrep	I	Neg emo	Pos emo	Cog mech	Ingest	Latency to 2 nd Entry	Word Count	Social Words	Com Rec'd	Diff Com	WC Com	Negem o Com	Posem o Com	Com to Others	Others Com To
Start Weight	.91***	.00	.01	.06	.05	-.13**	-.03	.15**	.16**	.16***	.07	.03	-.06	.01	.06	.06
Goal Discrepancy		.01	.07	.05	.05	-.13*	-.00	.03	.17***	.16*	-.06	-.00	-.05	.01	.10*	.10
I			.15**	.05	.26***	-.37***	.08	-.05	.08	.08	.05	.03	.05	-.01	.05	.04
Negemo				-.13**	.08	-.10*	.01	-.00	-.08	.02	-.01	-.01	.21***	-.17**	-.00	-.01
Posemo					.23***	-.26***	.00	-.15**	.32***	.12*	.12	-.02	-.04	.22***	.05	.05
Cogmech						-.43***	.08	.23***	.25***	.16***	.11*	.10*	-.02	-.19***	.10*	.10*
Ingestion words							-.01	-.19***	-.38***	-.15**	-.07	-.04	.01	.01	-.09	-.08
Latency to 2 nd entry								-.07	-.09	-.04	-.03	-.03	.02	.06	.01	.00
Total word count									.22***	.17***	.12*	.12*	-.03	-.18***	.09	.09*
Social words in blogs										.22***	.07	.02	-.09	-.02	.12*	.11*
Total comments in sample											.40***	.17***	.03	.06	.31***	.31***
#Different commenters												.95***	.08	-.01	.23***	.21***
Word count of comments													.08	-.11*	.20***	.19***
Negemo in comments														-.25***	.02	.01
Posemo in comments															.03	.04
#Comments to other blogs																.99***

Note. Pearson correlations. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$. Empty cells are redundant.

Appendix R: Correlations between Target Variables within the First Two Entries by Dedicated Bloggers

	Goal Discrep	I	Neg emo	Pos emo	Cog mech	Ingest	Latency to 2 nd Entry	Word Count	Social Words	Com Rec'd	Diff Com	WC Com	Negem o Com	Posem o Com	Com to Others	Others Com To
Start Weight	.91***	.17*	-.03	.16*	.05	-.02	.12	.00	.07	.12	.14	.16*	.15	.06	-.08	-.06
Goal Discrepancy		.15	-.05	.14	.04	.03	.15	-.03	.07	.06	.08	.09	.16	.10	-.07	-.05
I			.09	.26***	.23**	-.21**	.01	-.18*	-.04	-.04	.01	-.03	-.01	.08	-.15	-.12
Negemo				-.20**	-.12	.04	.11	.02	-.09	-.04	.01	.11	.11	-.14	-.05	-.06
Posemo					.23**	-.22**	-.02	-.27***	.30***	.08	.05	-.05	.12	.18*	.09	.11
Cogmech						-.26***	-.02	.13	.21**	.07	.04	.13	.07	-.09	.10	.10
Ingestion words							-.10	-.15	-.44***	-.09	-.05	-.04	.00	-.17*	-.01	.01
Latency to 2 nd entry								-.14	-.12	-.16*	-.14	-.12	.14	-.08	-.08	-.08
Total word count									.20**	.21**	.20**	.31***	.07	-.09	.01	-.00
Social words in blogs										.31***	.25***	.30***	.15	.04	.18*	.16*
Total comments in sample											.91***	.76***	.05	.06	.41***	.36***
Word count of comments												.77***	-.02	.07	.34***	.31***
Negemo in comments													.20*	-.31***	.33***	.30***
Posemo in comments														-.30***	-.00	-.02
#Comments to other blogs															-.03	-.03
																.98***

Note. Pearson correlations. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$. Empty cells are redundant.

Appendix S: Correlations of Target Variables during the First Two Entries with Group Membership (Dropouts = 0; Dedicated Bloggers = 1)

Dependent Variables	<i>r</i> with Group Membership	<i>p</i>	<i>df</i>
Start Weight	.08	.059	621
Goal Discrepancy	.07	.098	538
Self-Focus			
First person singular pronouns	-.08	.043	605
Emotions			
Negative emotion words	-.11	.008	605
Positive emotion words	-.04	.342	605
Cognitive Mechanisms			
Cognitive mechanism words	.00	.961	605
Self-Monitoring			
Focus on Food	-.08	.050	605
Ingestion words			
Rate of Blogging			
Latency to second entry	-.06	.141	621
Total word count	.19	.000	621
Social Support			
Social words in blogs	.08	.041	605
Total comments in sample	.06	.117	621
#Different commenters per blog	-.01	.872	621
Total word count of comments received	-.01	.881	621
Negative emotion word use in comments received	-.11	.015	504
Positive emotion word use in comments received	.02	.649	504
#Comments to other blogs	.25	.000	621
#Different blogs commented to	.25	.000	621

Note. Positive correlations indicate that the more a person was high on a target variable, the more likely they were to be a dedicated bloggers. Negative correlations indicate that the more a person was high on a target variable, the more likely they were to be a dropout.

Appendix T: Correlations of Percent Body Weight Loss over the 15 Weeks with all LIWC Categories, other Blogging Rates, and Comment Counts Measured within the First Two Entries of Dedicated Bloggers

		<i>M</i>	<i>SD</i>	<i>r</i> with % Weight Loss	<i>p</i>	<i>df</i>
Blog	LIWC CATEGORIES					
Entries	Linguistic Processes					
	Words/sentence	19.12	21.88	-.02	.814	174
	Dictionary words	88.58	5.76	.08	.321	174
	Words>6 letters	12.18	2.51	-.04	.604	174
	Numerals	2.38	2.32	.03	.674	174
	Total function words	57.73	5.59	.00	.980	174
	Total pronouns	17.12	3.25	.02	.778	174
	Personal pronouns	12.06	2.37	.02	.800	174
	1st pers singular	10.36	2.29	.13	.087	174
	1st pers plural	.36	.55	-.04	.569	174
	2nd person	.58	.74	-.15	.050	174
	3rd pers singular	.51	.87	-.08	.273	174
	3rd pers plural	.24	.34	-.13	.085	174
	Impersonal pronouns	5.06	1.59	.02	.843	174
	Articles	4.77	1.33	-.04	.641	174
	Common verbs	16.38	2.62	.04	.616	174
	Auxiliary verbs	10.17	2.06	.14	.067	174
	Past tense	3.83	1.73	-.08	.314	174
	Present tense	10.26	2.62	.08	.323	174
	Future tense	1.18	.83	.02	.754	174
	Adverbs	6.06	1.92	.05	.518	174
	Prepositions	12.37	1.79	-.16	.035	174
	Conjunctions	6.84	1.48	-.08	.307	174
	Negations	1.64	.80	-.14	.074	174
	Quantifiers	2.99	.97	.04	.601	174
	Numbers	.87	.54	.01	.886	174
	Swear words	.06	.17	-.11	.142	174
	Psychological Processes					
	Social processes	4.46	2.62	-.13	.093	174
	Family	.34	.46	-.11	.161	174
	Friends	.15	.23	-.01	.161	174
	Humans	.38	.42	.02	.805	174
	Affective processes	5.37	1.71	.15	.055	174
	Positive emotion	3.68	1.65	.11	.165	174
	Negative emotion	1.69	.81	.08	.276	174
	Anxiety	.24	.30	.08	.276	174
	Anger	.28	.38	-.01	.904	174
	Sadness	.77	.60	.15	.051	174
	Cognitive processes	15.55	2.63	-.05	.510	174
	Insight	1.88	.85	-.05	.556	174
	Causation	1.40	.75	.08	.291	174

Discrepancy	1.63	.92	-.01	.883	174
Tentative	2.59	1.00	.04	.651	174
Certainty	1.32	.75	.00	.977	174
Inhibition	.50	.44	-.08	.284	174
Inclusive	4.56	1.43	-.06	.472	174
Exclusive	2.87	1.14	-.08	.275	174
Perceptual processes	2.50	1.01	.04	.569	174
See	.65	.56	.06	.457	174
Hear	.27	.31	-.09	.262	174
Feel	1.29	.74	.03	.659	174
Biological processes	5.43	2.16	.09	.241	174
Body	.78	.73	.07	.377	174
Health	1.20	.81	.06	.430	174
Sexual	.16	.23	-.07	.357	174
Ingestion	3.84	2.01	.07	.376	174
Relativity	15.84	2.94	.01	.898	174
Motion	1.98	1.00	-.06	.429	174
Space	5.37	1.35	-.03	.738	174
Time	8.49	2.15	.07	.386	174
Personal Concerns					
Work	1.46	.88	.09	.253	174
Achievement	2.57	1.11	.21	.005	174
Leisure	1.40	.78	.02	.841	174
Home	.41	.56	-.20	.010	174
Money	.37	.42	-.09	.260	174
Religion	.09	.21	.18	.018	174
Death	.03	.08	-.11	.152	174
Spoken categories					
Assent	.26	.38	.09	.266	174
Nonfluencies	.41	.47	.11	.133	174
Fillers	.23	.28	-.02	.848	174
Comments	LIWC CATEGORIES				
Linguistic Processes					
Words/sentence	17.18	19.00	-.03	.767	146
Dictionary words	90.44	5.38	-.21	.011	146
Words>6 letters	12.31	2.84	.04	.629	146
Numerals	.86	1.00	.07	.413	146
Total function words	57.82	5.18	-.18	.033	146
Total pronouns	17.61	2.93	-.21	.012	146
Personal pronouns	12.29	2.49	-.23	.006	146
1st pers singular	5.15	2.20	-.08	.350	146
1st pers plural	.74	.75	-.03	.703	146
2nd person	5.70	2.29	-.14	.103	146
3rd pers singular	.34	.75	-.05	.517	146
3rd pers plural	.35	.47	-.04	.678	146
Impersonal pronouns	5.32	1.76	-.03	.757	146
Articles	5.36	1.46	.05	.520	146
Common verbs	16.92	2.87	-.01	.878	146
Auxiliary verbs	10.29	2.27	-.05	.561	146
Past tense	2.05	1.19	-.06	.468	146
Present tense	12.37	2.43	-.02	.86	146
Future tense	1.30	.76	.02	.828	146
Adverbs	5.94	1.89	-.12	.148	146

Prepositions	11.88	2.33	-.12	.135	146
Conjunctions	6.51	1.57	.09	.305	146
Negations	1.27	.87	-.05	.546	146
Quantifiers	2.83	1.37	.00	.961	146
Numbers	.66	.65	-.03	.687	146
Swear words	.03	.09	.21	.012	146
Psychological Processes			.00		
Social processes	11.69	3.50	-.15	.075	146
Family	.33	.65	-.07	.384	146
Friends	.14	.30	-.03	.734	146
Humans	.47	.57	-.04	.676	146
Affective processes	9.24	2.98	.00	.977	146
Positive emotion	7.94	3.15	-.01	.912	146
Negative emotion	1.39	.86	-.02	.801	146
Anxiety	.14	.25	.06	.513	146
Anger	.19	.30	-.03	.690	146
Sadness	.72	.65	.00	.965	146
Cognitive processes	16.46	2.96	-.11	.178	146
Insight	2.58	1.21	-.01	.906	146
Causation	1.32	.85	.01	.890	146
Discrepancy	1.70	1.01	-.09	.273	146
Tentative	2.97	1.43	-.02	.791	146
Certainty	1.37	.81	-.17	.043	146
Inhibition	.70	.64	.05	.544	146
Inclusive	4.78	1.57	.01	.947	146
Exclusive	2.64	1.36	-.09	.297	146
Perceptual processes	2.82	1.43	.07	.422	146
See	.92	1.06	.10	.212	146
Hear	.64	.66	-.11	.189	146
Feel	1.03	.72	.09	.271	146
Biological processes	3.44	1.73	.06	.441	146
Body	.44	.54	.06	.485	146
Health	.79	.70	.07	.424	146
Sexual	.33	.37	-.06	.504	146
Ingestion	2.09	1.43	.06	.483	146
Relativity	12.60	3.01	-.15	.066	146
Motion	1.67	.89	.05	.535	146
Space	5.43	1.47	-.12	.159	146
Time	6.03	2.23	-.17	.046	146
Personal Concerns			.00		
Work	1.74	1.11	-.05	.568	146
Achievement	2.92	1.40	-.09	.272	146
Leisure	1.30	.92	.01	.918	146
Home	.22	.35	-.14	.082	146
Money	.31	.62	-.05	.585	146
Religion	.18	.55	.05	.592	146
Death	.03	.09	-.01	.880	146
Spoken categories			.00		
Assent	.43	.59	-.01	.892	146
Nonfluencies	.29	.48	-.08	.316	146
Fillers	.45	.57	-.09	.295	146

Note. * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, two-tailed

Appendix U: Original Regression for Research Question 3 (Predicting Weight Loss in the Dedicated Bloggers based on the First Two Entries) with All Cases (i.e. including Outliers)

All of the target variables were entered into a stepwise regression equation. The regression equation was significant, $F(2, 173) = 17.87, p = .000$, and accounted for 17.12% of the variance in percentage body weight loss. The regression equation was:

$$\text{Percent Body Weight Loss} = 4.18 + .05 (\text{goal discrepancy}) - .34 (\text{social})$$

Beta Weights and their Statistical Significance in a Regression Equation Predicting Percent Body Weight Loss

Dependent Variables	Unstandardized Beta	SE	Standardized Beta	<i>t</i>	<i>p</i>
Constant	4.18	.98		4.29	.000
Goal discrepancy	.05	.01	.40	5.69	.000
Social words	-.34	.16	.15	2.21	.028

Note. $F(2, 173) = 17.87, p = .000, R^2 = .17$. Positive betas indicate that the more a person was high on a dependent variable, the more weight loss. Negative betas indicate that the more a person was high on a dependent variable, the more weight gain.

However, in this analysis, there was a single case with a standardized residual of percentage body weight loss greater than 4 standard deviations above the mean (i.e. 4.04), and so this case was excluded in a similar analysis reported in the dissertation.

References

- Adams, S. H. (2002). *Communication under stress: Indicators of veracity and deception in written narratives*. Unpublished Doctoral Dissertation, Virginia Polytechnic Institute and State University.
- Adams, S. (2004). Statement analysis: beyond the words. *FBI Law Enforcement Bulletin*, 73(April), 22-23.
- Back, M. D., Schmukle, S. C., & Egloff, B. (2008). How extraverted is honey.bunny77@hotmail.de? Inferring personality from e-mail addresses. *Journal of Research in Personality*, 42, 1116-1122.
- Baker, L., Wagner, T. H., Singer, S., & Bundorf, M. K. (2003). Use of the Internet and E-mail for health care information. *Journal of the American Medical Association*, 289, 2400-2406.
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1-15.
- Beebe, D. W., Holmbeck, G. N., Albright, J. S., Noga, K., & Decastro, B. (1995). Identification of "binge-prone" women: An experimentally and psychometrically validated cluster analysis in a college population. *Addictive Behaviors*, 20, 451-462.
- Berger, M., Wagner, T. H., & Baker, L. C. (2005). Internet use and stigmatized illness. *Social Science & Medicine*, 61, 1821-1827.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Black, D. R., Gleser, L. J., & Kooyers, K. J. (1990). A meta-analytic evaluation of couples weight-loss programs. *Health Psychology*, 9, 330-347.
- Boals, A., & Klein, K. (2005). Word use in emotional narratives about failed romantic relationships and subsequent mental health. *Journal of Language and Social Psychology*, 24, 252-268.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conception of time. *Cognitive Psychology*, 43, 1-22.
- Bucci, W., & Maskit, B. (2007). Beneath the surface of the therapeutic interaction: The psychoanalytic method in modern dress. *Journal of the American Psychoanalytic Association*, 44, 1355-1397.

- Bucci, W., & Maskit, B. (2005). Building a weighted dictionary for referential activity. In Y. Qu, J. G. Shanahan, & J. Wiebe (Eds.), *Computing attitude and affect in text* (pp. 49–60). Dordrecht, The Netherlands: Springer.
- Campbell, R.S., & Pennebaker, J.W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60-65.
- Chang-Schneider, C. S., & Swann, W. B. Jr. (2009). *Wearing self-esteem like a flag: Conveying our high- and low- self-esteem to others*. Manuscript submitted for publication.
- Chung, C.K., & Pennebaker, J.W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343-359). New York: Psychology Press.
- Chung, C. K., & Pennebaker, J. W. (2008). Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality, 42*, 96-132.
- Chung, C. K., Rentfrow, P. J., & Pennebaker, J. W. (2009). [This I Believe: Validity of themes mapped across America using text analysis]. Unpublished data.
- Cialdini, R. B. (1993). *Influence: The psychology of persuasion*. New York: William Morrow & Company, Inc.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*, 687-93.
- Cools, J., Schotte, D. E., & McNally, R. J. (1992). Emotional arousal and overeating in restrained eaters. *Journal of Abnormal Psychology, 101*, 348-351.
- Davison, K. P., Pennebaker, J. W., & Dickerson, S. S. (2000). Who talks? The social psychology of illness support groups. *American Psychologist, 55*, 205-217.
- de Castro, J. M. (1994). Methodology, correlational analysis, and interpretation of diet diary records of the food and fluid intake of free-living humans. *Appetite, 23*, 179-192.
- De Souza, P., & Ciclitira, K. E. (2005). Men and dieting: A qualitative analysis. *Journal of Health Psychology, 10*, 793-804.
- DietDiaries (www.dietdiaries.com). (2007). Homepage: "Welcome to Diet Diaries.com". Retrieved November 30, 2007, from DietDiaries, <http://www.dietdiaries.com/public-diary-support-blogs.cfm>

- Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. Oxford, England: Academic Press.
- Economos, C. D., Hildebrandt, M. L., & Hyatt, R. R. (2008). College freshman stress and weight change: Differences by gender. *American Journal of Health Behavior, 32*, 16-25.
- Eldredge, K. L., Agras, W. S., & Arnow, B. (1994). The last supper: Emotional determinants of pretreatment weight fluctuations in obese binge eaters. *International Journal of Eating Disorders, 16*, 83-88.
- Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology, 94*, 334-346.
- Fellows, M.D. & Pennebaker, J.W. (2008). *Emotion language among preschool aged children and parents in natural and lab settings*. Manuscript submitted for publication.
- Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage Publications.
- Fontaine, R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, p. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*, 1050-1057.
- Fox, S., & Fallows, D. (2003, July 16). Internet health resources. Retrieved June 03, 2007, from http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf
- Fredrickson, B. L. (2005). The broaden-and-build theory of positive emotions. In F. A. Huppert, N. Baylis, B. Kerverne (Eds.), *The science of well-being* (pp. 217-238). New York, NY: Oxford University Press.
- Gorin, A., Phelan, S., Tate, D., Sherwood, N., Jeffery, R., & Wing, R. (2005). Involving support partners in obesity treatment. *Journal of Consulting and Clinical Psychology, 73*, 341-343.
- Gorin, A. A., Phelan, S., Wing, R. R., & Hill, J. O. (2004). Promoting long-term weight control: does dieting consistency matter? *International Journal of Obesity, 28*, 278-281.
- Gottman, J. M., & Levenson, R. W. (2000). The timing of divorce: Predicting when a couple will divorce over a 14-year period. *Journal of Marriage & the Family, 62*, 737-745.

- Gottschalk, L. A. (2000). The application of computerized content analysis of natural language in psychotherapy research now and in the future. *American Journal of Psychotherapy*, 54, 305-311.
- Gottschalk, L. A., Winget, C. N., & Gleser, G. C. (1969). *Manual of instructions for using the Gottschalk-Gleser-Content Analysis Scales: Anxiety, Hostility, and Social Alienation-Personal Disorganization*. Oxford, England: University of California Press.
- Harmon, A. (2003). Finding comfort in strangers with an online diet journal. *The New York Times*. Retrieved October 7, 2007
<http://query.nytimes.com/gst/fullpage.html?res=9A06E2DD1239F936A1575BC0A9659C8B63>
- Heatherton, T. E., & Baumeister, R. F. (1991). Binge eating as escape from self-awareness. *Psychological Bulletin*, 110, 86-108.
- Heatherton, T. F., & Baumeister, R. F. (1996). Self-regulation failure: Past, present, and future. *Psychological Inquiry*, 7, 90-98.
- Heatherton, T. F., & Nichols, P. A. (1994). Personal accounts of successful versus failed attempts at life change. *Personality and Social Psychology Bulletin*, 20, 664-675.
- Heatherton, T. F., Striipe, M., & Wittenberg, L. (1998). Emotional distress and disinhibited eating: The role of self. *Personality and Social Psychology Bulletin*, 24, 301-313.
- Helvig, T. M., Sobell, L. C., Sobell, M. B., & Simco, E. R. (2006). Smoker's narrative accounts of quit attempts: Aids and impediments to success. *Psychology of Addictive Behaviors*, 20, 219-224.
- Henderson, M. D., Gollwitzer, P. M., & Oettingen, G. (2007). Implementation intentions and disengagement from a failing course of action. *Journal of Behavioral Decision Making*, 20, 81-102.
- Herman, C. P., & Polivy, J. (2003). Realistic and unrealistic self-change efforts. *American Psychologist*, 58, 823-824.
- Heshka, S., Anderson, J. W., Atkinson, R. L., Greenway, F. L., Hill, J. O., Phinney, S. D., Kolotkin, R. L., Miller-Kovach, K., Pi-Sunyer, F. X. (2003). Weight loss with self-help compared with a structured commercial program: A randomized trial. *Journal of the American Medical Association*, 289, 1792-1798.
- Hollis, J. F., Guillon, C. M., Stevens, V. J., Brantley, P. J., Appel, L. J., Ard, J. D., champagne, C. M., Dalcin, A., Erlinger, T. P., Funk, K., Laferriere, D., Lin, P.-H.,

- Loria, C. M., Samuel-Hodge, C., Vollmer, W. M., & Svetkey, L. P. (2008). Weight loss during the intensive intervention phase of the weight-loss maintenance trial. *American Journal of Preventive Medicine, 35*, 118-126.
- Ireland, M. E., & Pennebaker, J. W. (2009). *Language style matching in reading and writing*. Manuscript submitted for publication.
- Junghaenel, D. U., Smyth, J. M., & Santner, L. (2008). Linguistic dimensions of psychopathology: A quantitative analysis. *Journal of Social and Clinical Psychology, 27*, 36-55.
- Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2009). *The language of social hierarchies*. Manuscript submitted for publication.
- Kiesler, C. A., Pallak, M. S., & Kanouse, D. E. (1968). Interactive effects of commitment and dissonance. *Journal of Personality and Social Psychology, 8*, 331-338.
- Klein, K., & Boals, A. (2001a). Expressive writing can increase working memory capacity. *Journal of Experimental Psychology, 130*, 520-533.
- Klein, K., & Boals, A. (2001b). The relationship of life event stress and working memory capacity. *Applied Cognitive Psychology, 15*, 565-579.
- Koppel, M., Argamon, S., & Shimoni, A. (2003). Automatically categorizing written texts by author gender. *Literary and Linguistic Computing, 17*, 401-412 .
- Kuczmarski, M. F., Kuczmarski, R. J., & Najjar, M. (2001). Effects of age on validity of self-reported height, weight, and body mass index: findings from the Third National Health and Nutrition Examination Survey, 1988-1994. *Journal of the American Dietetic Association, 101*, 28-34.
- Legg, C., Puri, A., Thomas, N. (2000). Dietary restraint and self-reported meal sizes: Diary studies with differentially informed consent. *Appetite, 34*, 235-243.
- Leland, J. (2007). Debtors search for discipline via blogs. The New York Times. Retrieved October 7, 2007. <http://www.nytimes.com/2007/02/18/us/18debt.html>
- Lewinsohn, P. M., & Talkington, J. (1979). Studies on the measurement of unpleasant events and relations with depression. *Applied Psychological Measurement, 3*, 83-101.
- Low, C. A., Stanton, A. L., & Danoff-Burg, S. (2006). Expressive disclosure and benefit-finding among breast cancer patients: Mechanisms for positive health effects. *Health Psychology, 25*, 181-189.

- Lyons, E. J., Mehl, M. R., & Pennebaker, J. W. (2006). Pro-anorexics and recovering anorexics differ in their linguistic Internet self-presentation. *Journal of Psychosomatic Research, 60*, 253-256.
- Madara, E. J., Kalafat, J., & Miller, B. N. (1988). The computerized self-help clearinghouse: Using 'high tech' to promote 'high touch' support networks. *Computers in Human Services, 3*, 39-54.
- Marlatt, G. A., & Kaplan, B. E. (1972). Self-initiated attempts to change behavior: A study of New Year's resolutions. *Psychological Reports, 30*, 123-131.
- Martindale, C. (1990). *The clockwork muse: The predictability of artistic change*. New York, NY: Basic Books.
- McKenna, K. Y. A., & Bargh, J. A. (2002). Relationship formation on the Internet: What's the big attraction? *Journal of Social Issues, 58*, 9-31.
- Mehl, M. R. (2006a). Quantitative text analysis. In M. Eid and E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 141-156). Washington DC, USA: American Psychological Association.
- Mehl, M. R. (2006b). The lay assessment of subclinical depression in daily life. *Psychological Assessment, 18*, 340-345.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*, 862-877.
- Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology, 64*, 1306-1315.
- Miller, W. C., Koceja, D. M., & Hamilton, E. J. (1997). A meta-analysis of the past 25 years of weight loss research using diet, exercise or diet plus exercise intervention. *International Journal of Obesity Related Metabolic Disorders, 21*, 941-947.
- Mishne, G., & Glance, N. (2006). *Leave a reply: An analysis of weblog comments*. Paper presented at the 2006 Workshop on Weblogging Ecosystem, Edinburgh, UK.
- Mohler, P. P., & Zull, C. (1998). *Textpack (Windows 95/NT)*. Mannheim, Germany: ZUMA.
- Mullen, B., Chapman, J. G., & Peaugh, S. (1989). Focus of attention in groups: A self-attention perspective. *Journal of Social Psychology, 129*, 807-817.

- Nardi, B., Schiano, D., & Gumbrecht, M. (2004). Blogging as a social activity, or, would you let 900 million people read your diary? *Proceedings of Computer Supported Cooperative Work 2004*.
- National Public Radio. (2007, June 29). Blogging away debt? It can happen. Morning Edition Podcast retrieved from <http://www.npr.org/templates/story/story.php?storyId=11596225>
- Nau, R. F. (2009). Decision 411 Forecasting: Stepwise and all-possible-regressions. Webpage retrieved August 3, 2009, from <http://web.duke.edu/~rnau/regstep.htm>
- Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes, 45*, 211-246.
- Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*, 337-360.
- Norcross, J. C., Ratzin, A. C., & Payne, D. (1989). Ringing in the New Year: The change processes and reported outcomes of resolutions. *Addictive Behaviors, 14*, 205-212.
- Norcross, J. C., & Vangarelli, D. J. (1989). The resolution solution: Longitudinal examination of New Year's change attempts. *Journal of Substance Abuse, 1*, 127-134.
- Oberlander, J., & Gill, A. J. (2006). Language with character: A stratified corpus comparison of individual differences in E-mail communication. *Discourse Processes, 42*, 239-270.
- Odom, S. D. (2006). *A qualitative and linguistic analysis of an authority issues training group*. Unpublished doctoral dissertation, The University of Texas at Austin, Austin, TX.
- Oxman, T. E., Rosenberg, S. D., Schnurr, P. P., & Tucker, G. J. (1982). Diagnostic classification through content analysis of patients' speech. *American Journal of Psychiatry, 145*, 464-468.
- Pennebaker, J.W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science, 8*, 162-166.
- Pennebaker, J. W. (2009). [Computerized text analysis of George W. Bush's natural language use]. Unpublished data.

- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count (LIWC2007)*. Austin, TX: www.liwc.net.
- Pennebaker, J. W., & Chung, C. K. (2007). Expressive writing, emotional upheavals, and health. In H. S. Friedman and R. C. Silver (Eds.), *Foundations of health psychology* (pp. 263-284). New York, NY: Oxford University Press.
- Pennebaker, J.W., & Chung, C.K. (2008). Computerized text analysis of Al-Qaeda transcripts. In K. Krippendorff & M.A. Bock (Eds.), *A content analysis reader* (pp. 453-465). Thousand Oaks, CA: Sage.
- Pennebaker, J.W., Groom, C.J., Loew, D., & Dabbs, J. (2004). Testosterone as a social inhibitor: Two case studies of the effect of testosterone treatment on language. *Journal of Abnormal Psychology, 113*, 172-175.
- Pennebaker, J.W., & Harber, K.D. (1993). A social stage model of collective coping: The Loma Prieta Earthquake and the Persian Gulf War. *Journal of Social Issues, 49*, 125-145.
- Pennebaker, J. W., & Huddle, D. (2008). [Detecting deception with courtroom transcripts]. Unpublished data.
- Pennebaker, J. W. & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*, 1296-1312.
- Pennebaker, J.W. & Lay, T.C. (2002). Language use and personality during crises: Analyses of Mayor Rudolph Giuliani's press conferences. *Journal of Research in Personality, 36*, 271-282.
- Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*, 863-871.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547-577.
- Pennebaker, J.W. & Stone, L.D. (2003). Words of wisdom: Language use over the lifespan. *Journal of Personality and Social Psychology, 85*, 291-301.
- Perry, G. S., Byers, T. E., Mokdad, A. H., Serdula, M. K., & Williamson, D. F. (1995). The validity of self-reports of past body weights by U.S. adults. *Epidemiology, 6*, 61-66.

- Petrie, K.J., Pennebaker, J.W., & Sivertsen, B. (2008). Things we said today: A linguistic analysis of the Beatles. *Psychology of Aesthetics, Creativity, and the Arts*, 2, 197-202.
- Polivy, J., & Herman, C. P. (1985). Dieting and bingeing: A causal analysis. *American Psychologist*, 40, 193-201.
- Polivy, J., & Herman, C. P. (2002). If at first you don't succeed: False hopes of self-change. *American Psychologist*, 57, 677-689.
- Polivy, J., & Herman, C. P. (1999). The effects of resolving to diet on restrained and unrestrained eaters: The "false hope syndrome." *International Journal of Eating Disorders*, 26, 434-447.
- Presnell, K., Pells, J., Stout, A., & Musante, G. (2008). Sex differences in the relation of weight loss self-efficacy, binge eating, and depressive symptoms to weight loss success in a residential obesity treatment program. *Eating Behaviors*, 9, 170-180.
- Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47, 1102-1114.
- Prochaska, J. O., Norcross, J. C., & DiClemente, C. C. (1995). *Changing for good*. New York: Avon.
- Ramirez-Esparza, N., Chung, C. K., Kacwicz, E., & Pennebaker, J. W. (2008). The psychology of word use in depression forums in English and in Spanish. Testing two text analytic approaches. *Proceedings of the 2008 International Conference on Weblogs and Social Media*, pp.102-108.
- Ramirez-Esparza, N., Chung, C. K., Siera-Oterro, G., & Pennebaker, J. W. (2009). *Cross-constructions of self: American and Mexican college students*. Manuscript submitted for publication.
- Resnick, B., Orwig, D., Magaziner, J., & Wynne, C. (2002). The effect of social support on exercise behavior in older adults. *Clinical Nursing Research*, 11, 52-70.
- Rochon, E., & Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, 72, 193-218.
- Rude, S.S., Gortner, E.M., & Pennebaker, J.W. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion*, 18, 1121-1133.

- Ruderman, A. J. (1985). Dysphoric mood and overeating: A test of restraint theory's disinhibition hypothesis. *Journal of Abnormal Psychology, 94*, 78–85.
- Schupak-Neuberg, E., & Nemeroff, C. J. (1993). Disturbances in identity and self-regulation in bulimia nervosa: Implications for a metaphorical perspective of 'body as self.' *International Journal of Eating Disorders, 13*, 335-347.
- Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communications: The links between language, performance, error, and workload. *Human Performance in Extreme Environments, 5*, 63-68.
- Shiu, E., & Lenhart, A. (2004). How Americans use instant messaging. Retrieved October 14, 2005, from http://www.pewinternet.org/pdfs/PIP_Instantmessage_Report.pdf
- Sillars, A., Shellen, W., McIntosh, A., & Pomegranate, M. (1997). Relational characteristics of language: Elaboration and differentiation in marital conversations. *Western Journal of Communication, 61*, 403-422.
- Simmons, R. A., Gordon, P. C., & Chambless, D. L. (2005). Pronouns in marital interaction: What do 'you' and 'I' say about marital health? *Psychological Science, 16*, 932-936.
- Slatcher, R.B. & Pennebaker, J.W. (2006). How do I love thee? Let me count the words: The social effects of expressive writing. *Psychological Science, 17*, 660-664.
- Slobin, D. (1996). From “thought” and “language” to “thinking” for “speaking”. From J. J. Gumperz and S. J. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). New York, NY: Cambridge University Press.
- Smith, A. (2004). From words to action: Exploring the relationship between a group's value references and its likelihood of engaging in terrorism. *Studies in Conflict and Terrorism, 27*, 409-437.
- Sobell, L. C., Klingemann, H. K.-H., Toneatto, T., Sobell, M. B., Agrawal, S., & Leo, G. I. (2001). Alcohol and drug abusers' perceived reasons for self-change in Canada and Switzerland: Computer-assisted content analysis. *Substance Use & Misuse, 36*, 1467-1500.
- Steiger, H., Gauvin, L., Engelberg, M. J., Kin, N. M. K. N. Y., Israel, M., Wonderlich, S., A., & Richardson, J. (2005). Mood- and restraint-based antecedents to binge episodes in bulimia nervosa: possible influences of the serotonin system. *Psychological Medicine, 35*, 1553-1562.

- Stirman, S.W., & Pennebaker, J.W. (2001). Word use in the poetry of suicidal and non-suicidal poets. *Psychosomatic Medicine*, *63*, 517-522.
- Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). *The General Inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.
- Stone, L.D. & Pennebaker, J.W. (2002). Trauma in real time: Talking and avoiding online conversations about the death of Princess Diana. *Basic and Applied Social Psychology*, *24*, 172-182.
- Subrahmanyam, K. Smahel, D., & Greenfield, p. (2006). Connecting developmental constructions to the Internet: Identity presentation and sexual exploration in online teen chat rooms. *Developmental Psychology*, *42*, 395-406.
- Taylor, S. E. (2007). Social support. In H. S. Friedman and R. C. Silver (Eds.), *Foundations of health psychology* (pp. 145-171). New York, NY: Oxford University Press.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403-421.
- Tsai, J. L., Simeonova, D. I., & Watanabe, J. T. (2004). Somatic and social: Chinese Americans talk about emotion. *Personality and Social Psychology Bulletin*, *30*, 1226-1238.
- Tsai, A. G., & Wadden, T. A. (2005). Evaluation of the major commercial weight loss programs. *Annals of Internal Medicine*, *142*, 56-66.
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, *19*, 95-112.
- Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, *87*, 123-132.
- Vigas, F. B. (2005). Bloggers' expectations of privacy and accountability: An initial survey. *Journal of Computer-Mediated Communication*, *10*, article 12.
- Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science*, *11*, 249-254.
- Wallis, D. J., & Hetherington, M. M. (2004). Stress and eating: The effects of ego-threat and cognitive demand on food intake in restrained and emotional eaters. *Appetite*, *43*, 39-46.

- Weinstein, P. K. (2006). A review of weight loss programs delivered via the Internet. *Journal of Cardiovascular Nursing, 21*, 251-258.
- Weintraub, W. (1989). *Verbal behavior in everyday life*. NY: Springer.
- Wheeler, L., & Reis, H. T. (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of Personality, 59*, 339-354.
- Williams, A. (2006). Before spring break, the anorexic challenge. The New York Times. Retrieved October 7, 2007.
<http://www.nytimes.com/2006/04/02/fashion/sundaystyles/02BREAK.html>
- Wing, R. R., & Hill, J. O. (2001). Successful weight loss maintenance. *Annual Review of Nutrition, 21*, 323-341.
- Wing, R. R., & Jeffery, R. W. (1999). Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. *Journal of Consulting and Clinical Psychology, 67*, 132-138.
- Wolf, M., Chung, C. K., & Kordy, H. (in press). Inpatient treatment to online aftercare: E-mailing themes as a function of therapeutic outcomes. *Psychotherapy Research*.
- Womble, L. G., Wang, S. S., & Wadden, T. A. (2002). Commercial and self-help weight loss programs. In T. A. Wadden (Ed.), *Handbook of obesity treatment* (pp.395-415). New York, NY: Guilford Press.
- Zapata, C. (2007). Dieters finding willpower on Web. San Antonio's Express-News. Retrieved October 7, 2007.
<http://www.mysanantonio.com/salife/columnists/czapata/stories/MYSA061107.1P.zapata.25244ea.html>

Vita

Cindy Kyuah Chung was born in Toronto, Ontario, Canada on October 20, 1978 to 정경문 and 김종금. Cindy, her older sister, Lois S. Chung, and younger brother, A. Jinu Chung all attended The University of Toronto, where Cindy earned her bachelor's degree in Psychology in 2002. She worked for two years at the University of Toronto for Professor C. Peter Herman and Lenny R. Vartanian. Beginning in August of 2003, she studied under Professor James W. Pennebaker in the Social and Personality Psychology doctoral program at The University of Texas at Austin. During that time, Cindy was a consultant and senior investigator for various text analysis projects in the counterintelligence field. She has been involved in the development of text analysis tools and methods such as the Meaning Extraction Method, Linguistic Inquiry and Word Count 2007, and Arabic Linguistic Inquiry and Word Count. She has published her work on psychological assessments based on natural language in multiple journal articles and book chapters. Upon graduation, she will be assuming a post-doctoral position with Professor Pennebaker at the University of Texas at Austin.

Permanent address: 236 Rizal Avenue, Markham, ON, CANADA L6B 0G5

This dissertation was typed by the author.