Seasonal and Event-related Fluctuations in the Positivity of the Language in Sir Walter Scott's Journal

Cynthia Whissell
Psychology Department,
Laurentian University.

ABSTRACT
Computerized sentiment analysis based on the Dictionary of Affect in Language (Whissell, 2009) was employed to assess patterns in the employment of positive (pleasant) language in the journal kept by Sir Walter Scott between 1825 and 1832. The continuous portion of the journal (33 months) was modeled with a polynomial regression. Two major patterns were observed. In the earlier portions of the journal, significant seasonal variations were present, with more positively toned entries in the Spring/Summer and less positive ones in the Fall/Winter. In the later portions of the journal there were discontinuities most likely attributable to ill health; significantly greater emotional variability was noted with some of the least positive entries occurring at about the time when Scott was facing his final illness and just before he stopped making journal entries.

SEASONAL AND EVENT-RELATED FLUCTUATIONS IN THE POSITIVITY OF THE LANGUAGE IN SIR WALTER SCOTT'S JOURNAL

Sir Walter Scott (1771-1832), the famous author, kept a journal during the last seven years of his life (Scott, 1891). Scott fully expected this journal to be published after his death: he referred to "the public" as a potential audience in his very first entry of November 20th, 1825. Every month had several entries associated with it between November of 1825 and July of 1828, but after this point (up to April, 1832) the diary was maintained for segments of several months with breaks interspersed among them. Scott was still actively writing for publication at this time. His journal entries were permeated with two main concerns - his ongoing physical ailments and his efforts to successfully discharge a bankruptcy that was declared in early 1826. This article examines the positivity of the language in Scott's journal. Language was scored by a computer program employing the Dictionary of Affect in Language (Whissell, 2009) which assigns pleasantness values, based on the opinions of independent raters, to words. It was expected that fluctuations in positivity, represented by the pleasantness of Scott's language, would be related to events in his life. The employment of emotion in diary entries to predict and explain behavior has been attempted by researchers such as Lester (2009) who used LIWC computerized sentiment analysis of diary contents to examine the emotions of an author immediately previous to the author's suicide. Data from Scott's journal were a rich source of information. As described further below, they fulfilled the expectation that positivity would be related to life events and also indicated the presence of distinct seasonal variations in positivity as well as extreme fluctuations in positivity within the latter – discontinuous – portion of the journal.

During the time period represented by the journal, Scott encountered several situations with the potential to influence the positivity of his writing; some of these are outlined in Table 1. Potentially impactful events were associated with his bankruptcy. Scott was not parsimonious, and he was generous with his family. However, his bankruptcy was probably the result of the market crash of November, 1825 (the month he began his journal), and the fact that Scott had

guaranteed the loans of others (McKinstry & Fletcher, 2002), rather than being the outcome of his spending habits. Scott worked for the rest of his life to pay off all his creditors, but this was only accomplished after his death with the help of his literary estate. Scott's health was also an ongoing concern: as well as suffering a series of strokes later in life (Table 1), Scott was subject to various other repetitive ailments such as headaches, pains in the legs, and lameness (Oxford Dictionary of Biography; n.d.).

The presence of seasonal variations in positivity was not an *a priori* prediction of the research, but was definitely characteristic of the data discussed below. The extreme form of systematic seasonal variations in mood is referred to as Seasonal Affective Disorder. This "was originally defined as a syndrome in which depression developed during autumn or winter and remitted in the following spring or summer" (Magnusson & Partonen, 2005). Similar but less extreme systematic fluctuations in mood are referred to as "subsyndromal seasonal affective disorder" (Magnusson & Partonen, 2005) or simply "mood rhythms" (Golder & Macy, 2011). Golder and Macy (2011) report that a less positive mood is associated with seasons of shortening days. The presence of seasonal changes in positivity for Scott's journal would be confirmed by fluctuations in affect with greater positivity being apparent during spring and summer months and less positivity during fall and winter months.

METHOD
Scott's journal was downloaded in plain text from Project Gutenberg.[1] It contained some 258 thousand words. All words were compared by a computer program to the list for the Dictionary of Affect, and those found in the Dictionary (225,368 words, or 87.5% of the total) were assigned scores for pleasantness. The pleasantness of individual words had been determined, when the Dictionary was originally established, by asking raters to rate the implications or associations of a word on the scale of unpleasant-pleasant. Average pleasantness was calculated for each diary entry (by day), and for each month where data were available. Word pleasantness represented positivity in journal entries. Words such as "alarmed", "grim", "grief", "against", "fatigue", "fear", "inability", "pain", "poor", "sadly", "uncomfortable", and "waste" had extremely low Dictionary scores for pleasantness and contributed to lower overall positivity. Words such as "accomplish", "extraordinary", "friendship", "fun", "laugh", "profit", "satisfied", "smile", "wealth", and "wisdom" had extremely high scores for pleasantness and contributed to higher positivity.

As a caution, it should be noted that Dictionary of Affect scoring was applied to a form of English in common use more than 150 years before the Dictionary itself was created. The pleasantness or positivity of some individual words will likely have changed across time. However, there was a substantial scoring rate for the Dictionary, words such as those exemplified above still have similar emotional connotations, and the Dictionary has been employed successfully with even older materials such as Shakespeare's dramas (Whissell, 2008). The age of the materials studied is a source of valid but not excessive concern unless the differences it contributes are systematic in some way, which is unlikely. Statistically significant and useful information can be evaluated appropriately in spite of the age of the documents.

RESULTS AND DISCUSSION
Modeling Continuous Data
Since relatively continuous data were available for the early years, the pleasantness of 928 daily entries from November, 1825 to July, 1828 was modeled with a full entry polynomial

---

[1] https://www.gutenberg.org/ebooks/14860

regression including standardized elapsed time and the second to seventh powers of standardized time. Terms up to the 12th power were essayed without additional benefits to prediction. The seven term model was significant ($p<.002$; $R=.15$), and all predictors involving an odd power of time were individually significant within it. During the continuous period, pleasantness rose and fell repeatedly. The predicted pleasantness mean for each of the 33 continuous months is plotted in the first half of Fig. 1. Months are represented on the horizontal axis. The correlation between actual monthly pleasantness and predicted monthly pleasantness was .73 ($p<.01$) in the continuous period, showing that predictions based on polynomial regression were a good fit at this level of analysis. The modeled plot is associated with several high points, all of them taking place in spring/summer, and several low points, all but the last taking place in fall/winter. The last drop within the continuous period comes at the end of the 33 months, where June and July exhibit atypically low pleasantness scores before the diary is discontinued. In his July 1828 entries made at the end of the period, Scott refers several times to headaches and also to sadness, which might explain the drop.

Discontinuous Data

Discontinuous data were not modeled, so the data plotted in the second half of Fig. 1 are raw monthly means. It is evident that Scott took up his journal several times, maintained it for a few months, and then discontinued his entries once more. These bursts of activity are labeled in Fig. 1 as A, B, C, and D. Each is characterized by a rise in positivity. Several points related to key events from Table 1 are represented directly on the graph by their number. Scott's bankruptcy is located at a local minimum for positivity (number 1). His first stroke took place in the time between A and B (number 6), his second stroke in the time between B and C (number 8), his third stroke towards the end of C (10) and his final and fourth stroke after he had permanently discontinued diary entries (after D; 11). The author's strokes were likely preceded by general ill health which may point to why the journal was discontinued. At the time that segment D was being written, Scott was travelling for his health in Europe and his health was deteriorating rather dramatically (Buchan, 1961; p. 328).

There was no significant difference between the monthly mean raw pleasantness of entries during the continuous and the discontinuous period (50.27, 50.26, $t_{56}=.06$, $p>.95$) but the variability of scores in the two period was different. Monthly raw means in the continuous period were much less variable (variance=.34) than those in the discontinuous period (variance=1.12; $F=6.06$, $p<.02$). This wide variation in the positivity of entries during the discontinuous period is evident in Fig. 1 where points fluctuate dramatically, reaching high and low extremes never reached during the continuous period, even in raw scores. The high points in A, B, and C occur during spring/summer, maintaining the seasonal pattern somewhat, while D shows a profound (and not unsurprising) decline in pleasantness during the summer period leading up to Scott's final stroke. When the monthly totals for pleasantness in the data set were compared for the fall-winter months (November, December, January and February) and the spring-summer months (May, June, July, August), the latter were found to be significantly more pleasant (50.5, 49.7; $t_{39}=2.09$, $p<.05$), although the difference was a small one ($r^2=.11$). In spite of the inconsistent emotional impact of events characterizing months within the two seasons, the overall difference was in the direction predicted by seasonal mood variations.

The Death of Scott's Wife

The low positivity of Scott's language at the beginning of the modeled curve in Fig. 1 could be attributed to the season (fall/winter), to his escalating monetary problems, or to some combination of these factors. It is puzzling, however, that the death of Scott's wife Charlotte, in May of 1826, occurs at one of the more positive points in the modeled curve. An examination of the raw positivity of entries for the day where Scott noted his wife's death (May 15th, 49.18, a

brief entry), and the day where he discussed it at some length (May 16th, 47.93) did actually evince relatively low positivity. Daily positivity was normally distributed with a mean of 50.35 and a standard deviation of 2.38. The z scores for the two days were low (-.45, -.90) but not among the lowest. The positivity for these dates was not even in the lowest decile of daily entries and the days were not outliers within the polynomial regression. The least positive daily entries (made in December 1825 and January 1826) were concerned with Scott's realization of his mounting financial difficulties and with problems associated with his writing under pressure. The lowest monthly averages characterized the months immediately before his death. The reason for the moderately rather than extremely low positivity of entries associated with Scott's wife's death probably lay in Scott's daily expectation of this event. In his entry for May 11th of 1826 he mentions "I have foreseen, for the last two years and more, that this menaced event could not be too far distant ... that recovery was hopeless." On May 16th, after her death, he wrote "... I suspect they [his thoughts] will be hers [his wife's] for a long time at least. But I will not blaze cambric and crape in the public eye like a disconsolate widower ..." These entries suggest that Scott was not so much unfeeling at his wife's death as he was prepared for it, which is often the case for those whose loved ones die after a lengthy illness.

CONCLUSIONS

Several conclusions can be reached on the basis of the data discussed here. The positivity of the language in Scott's journal entries evinced clear fluctuations in response to events in his life. Monetary pressures (e.g., number 1 in Fig. 1) were associated with low positivity, as were health challenges (e.g., number 11). Health challenges were also associated with breaks in the maintenance of the journal (e.g., numbers 6, 8) and greater emotional variability. Distinct seasonal variations with higher positivity in spring/summer and lower positivity in fall/winter were documented in the model of the continuous data and in the entire journal.

Egan (1978) suggests that the plot of a narrative may be sought in its systematic emotional variations across time. In this study, variations in the positivity of Scott's journal entries provide a "plot" for his life. They show the negative influence of various factors, and the positive influence of others. And they highlight bursts of effort towards the end of his life. The type of analysis performed here contributes to a better understanding of the author's later life, and of the period associated with and following his bankruptcy.

References

Buchan, J. (1961) *The Life of Sir Walter Scott*. 8th Edition Cassell & Co., London

Egan, K. (1978) What is a plot? New Literary History, 9, 455-473.

Golder, S. A. & Macy, M. W. (2011) Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333, 1878-1881.

Lester, D. (2009) Learning about suicide from the diary of Cesare Pavese. *Crisis*, 30, 222-224.

Magnussen, A. & Partonen, T. (2005) The Diagnosis, symptomatology, and epidemiology of seasonal affective disorder. *CNS Spectrums*, 10, 625-634.

McKinstry, S. & Fletcher, M. (2002) The personal account books of Sir Walter Scott. *Accounting Historians Journal*, 29, 59-89.

Oxford Dictionary of National Biography (n.d.) https://www.oxforddnb.com/view/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-24928?rskey=w48kON&result=1

Whissell, C. (2008). *A psychological investigation of the use of Shakespeare's emotional language: The case of his Roman tragedies*. Lewiston, NY: Edwin Mellen Press.

Whissell, C. (2009). Using the revised Dictionary of Affect in Language to quantify the emotional undertones of samples of natural language. *Psychological Reports*, *105*, 1-13.

Scott, W. (1891) *The journal of Sir Walter Scott 1825-1832*. D. Douglas, Ed. David Douglas, Edinburgh.

Table 1 Some impactful events encountered by Scott during the years when he kept a journal and the codes associated with these events in Fig. 1

Month/Year	Code in Fig. 1	Event
January, 1826	1	The chain of events leading to Scott's bankruptcy begins
May, 1826	2	Scott's wife dies after a lengthy illness
June, 1826	3	Scott's accounts come under control of the trust managing his bankruptcy
December, 1827	4	The trust established to pay off Scott's debts makes a first payment
July, 1829	5	Thomas Purdie, Scott's friend and "henchman", dies
February, 1830	6	Scott suffers his first stroke
July, 1830	7	The trust makes an additional payment
November, 1830	8	Scott suffers his second stroke
December, 1830	9	Scott recovers rights to his library and household possessions
April, 1831	10	Scott suffers his third stroke
April, 1832	11	Journal terminates.
May, 1832		Scott suffers a fourth stroke,
September, 1832		Scott dies

Sources: Buchan (1932); Scott (1891); Oxford Dictionary of National Biography (n.d.)

Figure 1. Fluctuations in pleasantness (positivity) across months for Scott's journal: numbers within the graph denote events listed in Table 1 and the letters A-D label the discontinuous segments of the journal