Journal of Personality 00:00, Month 2015

DOI: 10.1111/jopy.12183

A Comparison of Human Narrative Coding of Redemption and Automated Linguistic Analysis for Understanding Life Stories

Sara J. Weston,[1] Keith S. Cox,[2] David M. Condon,[3] and Joshua J. Jackson[1]
[1]Washington University in St. Louis
[2]Medical University of South Carolina
[3]Northwestern University

Abstract

The majority of life narrative research is performed using trained human coders. In contrast, automated linguistic analysis is oft employed in the study of verbal behaviors. These two methodological approaches are directly compared to determine the utility of automated linguistic analysis for the study of life narratives.

In a study of in-person interviews ($N = 158$) and a second study of life stories collected online ($N = 242$), redemption scores are compared to the output of the Linguistic Inquiry and Word Count (Pennebaker, Francis & Booth, 2001). Additionally, patterns of language are found using exploratory principal components analysis.

In both studies, redemption scores are modestly correlated with some LIWC categories and unassociated with the components. Patterns of language do not replicate across samples, indicating that the structure of language does not extend to a broader population. Redemption scores and linguistic components are independent predictors of life satisfaction up to 3 years later.

These studies converge on the finding that human-coded redemption and automated linguistic analysis are complementary and nonredundant methods of analyzing life narratives, and considerations for the study of life narratives are discussed.

INTRODUCTION

Three decades of theory and empirical findings support the contention that the *life story*—the narratives individuals construct to make sense of their personal past, lived present, and anticipated future—is a central aspect of human personality (e.g., Hammack, 2008; McAdams & Pals, 2006; McLean, Pasupathi, & Pals, 2007; Singer, 2004; McAdams, 1995; Josselson & Lieblich, 1993). Through the life story, individuals create a narrative identity that communicates who they understand themselves to be and who they think they will become. Numerous studies show that characteristics of life stories such as thematic content, valence dynamics, and forms of autobiographical reasoning are meaningfully associated with important life outcomes, even when controlling for other personality factors (Dunlop & Tracy, 2013a; Lodi-Smith, Geise, Roberts, & Robins, 2009; Manczak, Zapata-Gietl, & McAdams, 2014; McAdams, Reynolds, Lewis, Patten, & Bowman, 2001; McLean & Pratt, 2006). Still, researchers recognize the need for the science of the life story to mature (e.g., Adler, Lodi-Smith, Philippe, & Houle, 2015), as basic questions remain unaddressed.

One of the most fundamental questions is whether different methods of narrative analysis provide unique utility to the study of life stories. The primary method is to use trained human coders who analyze stories for the presence of constructs such as episodic coherence or autobiographical reasoning (e.g., Dunlop & Tracy, 2013b). Human narrative coding of life stories has been shown to predict or be associated with a diverse set of life outcomes and individual difference constructs, including well-being (Adler et al., 2015), recovery from alcoholism (Dunlop & Tracy, 2013a), Big Five personality traits (McAdams et al., 2004), emotion regulation (Cox & McAdams, 2014), political orientation (McAdams et al., 2008), and psychotherapy outcomes (Adler, 2012).

This study was supported by a grant from the Luxembourgish Fonds National de la Recherche (PELEDUPersonality in Learning and Education, project no. C11/LM/1168993).

Correspondence concerning this article should be addressed to Sara Weston, Department of Psychology, Washington University in St. Louis, Campus Box 1125, St. Louis, MO 63130. Email: sweston@wustl.edu.

In the study of life narratives, techniques other than human narrative coding are rarely used to analyze the spoken or written word. However, in contrast to human coding, language can be analyzed through automated linguistic analysis techniques, such as Linguistic Inquiry Word Count (LIWC: Pennebaker, Francis, & Booth, 2001). These automated linguistic analysis programs evaluate life narratives at the level of individual words, word categories (e.g., personal pronouns or causal words), and frequency of word use. The utility of automated linguistic analysis techniques can be seen in several studies of health and well-being. For example, word use predicts recovery from trauma (Pennebaker, Mayne & Francis, 1997), psychological adjustment to major health events (Robbins, Mehl, Smith, & Weihs, 2013), depression (Rude, Gortner & Pennebaker, 2004) and physical health (Eichstaedt et al., 2015; Kern et al., 2014; Pennebaker & King, 1999; Riley, Snowden, Derosiers, & Markesbery, 2005), as well as Big Five personality traits (Hirsh & Peterson, 2009).

There are numerous advantages of automated linguistic analysis over human coding. First, the level of analysis is at the level of individual words or word categories. This elemental analysis allows researchers to tease apart what participants are saying as opposed to the overall idea they mean to convey, with the latter better captured through a more thematic rating through human narrative coding. Second, automated linguistic analysis can be further analyzed to uncover patterns or structures of word use that may not be apparent at the level of individual words. For example, if individuals tend to use past tense verbs along with negative emotion words, they may be ruminating, or thinking about past negative events, with the outcome of prolonging negative emotions. Third, automated linguistic analysis is relatively efficient compared to labor-intensive human narrative coding.

Comparisons Between Human Narrative Coding and Automated Linguistic Analysis

Despite the promise of both techniques, a direct comparison of these two methods within life narrative research has not to our knowledge been investigated in the literature. In the only example of which we are aware that employs the two methods simultaneously, human narrative coding of storymaking was weakly associated with automated linguistic analysis in a study of essays (Graybeal, Sexton & Pennebaker, 2002). Currently it is unknown whether the coding methods used in the study of life stories can be captured through automated linguistic analyses or whether the two methods provide unique insights into the life story. In other words, do the linguistic properties of a life narrative exhaust the psychologically meaningful aspects of a narrative? Or do human coders pick up on aspects of the story that cannot be boiled down to the words people use to tell those stories?

Creating a life story is a verbal behavior, and thus we may expect that the coded constructs found in life narratives are strongly tied to the words used. If the two methods utilize the same content then life narrative researchers could potentially use automated linguistic analysis to more efficiently code constructs of interest. If the two are different, some aspects of life narratives can be thought of as emergent properties that may only be captured through more thematic methods of analysis, such as human narrative coding. Furthermore, if the two methods are different, then human narrative coding and automated linguistic analysis capture complementary aspects of life narratives. For example, human coders often rate the narrative arc of a story, such as whether it ends positively. Meanwhile, automated linguistic analysis can measure the complexity of sentences (Riley et al., 2005). Knowing that a sequence ends positively would not reveal whether the narrative has complex sentences; furthermore, narrators who use complex sentences would not be more or less likely to end their stories in positive ways.

A direct comparison would address the extent to which these methods overlap and uniquely predict relevant outcomes. We see three possible outcomes in such a comparison. One, human narrative coding and automated linguistic analysis yield highly corresponding results, i.e., the respective kinds of data are highly redundant, suggesting that human narrative coding does not capture unique variance beyond word frequencies and patterns of word use. This could count as strong evidence for the nonexistence of emergent properties of life narratives. Second, human narrative coding and automated linguistic analysis yield moderately corresponding results, i.e., the respective kinds of data are moderately associated, suggesting human narrative coding and automated linguistic analysis tap both overlapping and nonoverlapping features of life narratives. This would count as clear evidence for the existence of emergent properties of life narratives and necessitate the continued use of human narrative coding. Three, human narrative coding and automated linguistic analysis result in wholly noncorresponding results, i.e., the respective kinds of data are uncorrelated, suggesting that human narrative coding and automated linguistic analysis tap nonoverlapping aspects of life narratives. This would count as strong evidence that much of the results from human narrative coding are attributable to emergent properties of, as opposed to the word use in, life stories. Additionally, this would suggest that automated linguistic analysis offers important information currently untapped by coding. The current study compares human narrative coding and automated linguistic analysis with this theoretical framing in view.

Redemption as Test Case for Emergent Properties

In the current study, we select the construct of narrative redemption as the focus of study. Narrative redemption refers to a sequence in a life story episode in which the scene transitions from a fundamentally negative state to one including, at least in part, positive elements (McAdams, 2006; McAdams, Diamond, de St. Aubin, & Mansfield, 1997). In a redemptive narrative, the narrator may describe the transition from a negative state to a positive one with different kinds of explanations or rationales, such as believing that the negative event culminated in a positive event, or the negative event resulted in personal growth and learning. Redemption sequences are popular in American culture (McAdams, 2006) and are among the most widely assessed content of life narratives, in part because they are linked with higher levels of self-reported psychological well-being (Lilgendahl & McAdams, 2011; McAdams et al., 2001). It is theorized that telling a redemptive sequence helps individuals derive meaning from adverse events (Singer, 2004) and reconcile negative events with aspects of their identity (Adler & Poulin, 2009; McLean & Pratt, 2006), providing a positive framework for their lives (King, 2001).

There are several reasons redemption was chosen to compare the two methods. First, redemption has been widely studied (e.g., McAdams, 2006; Benish-Weisman et al., 2014; McLean & Breen, 2009), has been shown to relate to multiple aspects of personality (e.g., Lodi-Smith et al., 2009; McAdams & Guo, 2015), and predicts longitudinal outcomes (e.g., Dunlop & Tracy, 2013a). Thus, redemption sequences are one of the most commonly studied constructs by life narrative researchers. Second, the valence features of redemption (negative to positive) are likely to have a verbal signature that automated linguistic analysis is readily able to detect. In other words, human coders may be subconsciously picking up on verbal cues within a text that signify redemption, rather than thematically taking in the emergent properties. If so, then word use and patterns of word use might be enough to account for a seemingly complex narrative phenomenon like redemption. In this way, redemption provides a suitable first test case for considering the relative extent of word use and emergent properties of narratives.

The Current Study

The current study aims to determine the relationship between automated linguistic analysis and human narrative coding with regard to the narrative construct of redemption. To do so, two separate samples of language data will be analyzed. The first sample comes from a longitudinal study of aging adults who provided in-person narratives; the second sample was collected online. Narrative redemption will be studied in individuals' stories of adverse experiences, or low points, as these kinds of stories are thought to be the most likely places for redemptive sequences to appear. Moreover, as valenced word categories, like positive and negative emotion words, are likely to be present in stories of adverse, personal experiences, these kinds of stories will likely provide many targets for automated linguistic analysis of valenced words and word categories.

We sought to answer the following questions:

1. How do categories of language relate to the human narrative coding of redemption? Automated linguistic analysis in both samples will be compared with redemption scores generated by human coders to assess how these methods of narrative analysis are related. Given that we are trying to determine whether narratives can be understood as the sum of their parts, we will use the entire catalogue of LIWC categories to measure as many words as possible in these stories. Redemption sequences are defined by descriptions of both positive and negative content; consequently, it is expected that redemption scores will be positively related to both the *positive* and *negative emotion* categories.

2a. Are there patterns of word use in life narratives? Word use is a fairly complex behavior, and individual word categories may not be sufficient to fully understand a sample of linguistic data. We plan to use all LIWC categories but this will yield a large number of statistical tests that may be difficult to interpret individually. Furthermore, few or none of the categories may be associated with human narrative coding of redemption in isolation while combinations of words are. To organize and conceptualize the word categories, exploratory principal components analysis will be applied to the LIWC categories. Akin to using factor analysis to simplify the complexity of personality items to the Big Five traits, we use principal components to simplify the automated linguistic analysis and explore the possibility of components. The primary benefit of using principal components is the preservation of variance—we are able to reduce the number of associations we test while incorporating as much variability in word use as possible. To our knowledge, only two studies have examined patterns of language (Chung & Pennebaker, 2008; Pennebaker & King, 1999). Pennebaker and King (1999) found reliable and interpretable principle component structure for LIWC measures across time and topic, and this structure predicted real-world behaviors and outcomes, while Chung and Pennebaker (2008) expanded the measurement of words to include specific adjectives and found similar principal components. Word use in these studies came from stream-of-consciousness essays, daily diaries, scholarly papers, and self-descriptions. Since then, no study has replicated their methods or examined patterns of word use. Furthermore, no study has attempted to uncover the structure of word use in life stories. This study will examine the components of word use in low points to explore whether these components reflect meaningful individual differences in these stories.

2b. Does the structure of word categories relate to redemption scores? The principal components discussed above are hypothesized to reflect meaningful differences in low-point stories; one such difference may be the inclusion of redemptive sequences. Components will be compared both qualitatively (looking for similar characteristics in the word categories which comprise a linguistic component and the defining qualities of redemption sequences) and quantitatively.

3. Do redemption coding and automated linguistic analysis each provide unique utility to the study of life narratives? Regardless of whether redemption and word use are related, the unique utility of automated linguistic analysis has yet to be determined. That is, does automated linguistic analysis allow researchers to understand something additional from a narrative, beyond what is captured by typical human narrative coding? To investigate we looked at whether well-being is predicted by both redemption scores and word use, up to 3 years after the life story. Both redemption (McAdams et al., 2001) and word use (Pennebaker et al., 1997) predict well-being, but it is currently unknown whether these methods overlap and capture the same variability in well-being or whether each taps into a different component of well-being.

STUDY 1 METHODS

Participants

Participants ($N = 158$; 64% female) were recruited in the Chicago area through the Foley Longitudinal Study of Adulthood (Foley; see Cox & McAdams, 2014; Manczak et al., 2014), a study of personality development in late-midlife U.S. adults ($M_{age} = 56.03$ years, $SD_{age} = 4.69$). The sample is 56% White, 42% African American, and 2% Interracial/Other. The sample was well educated (70.9% completing high school, 46.2% earning at least a bachelor's degree) and had relatively high income (10.8% made less than $25,000/year, 36.7% made more than $75,000). Participants were invited to come back 3 years after baseline, at which time they completed the same self-report questionnaires as baseline ($N = 143$). Individuals who did not complete the follow-up did not differ on redemption scores ($M_{completed} = 0.31$, $SD_{completed} = 0.39$; $M_{not\ completed} = .20$, $SD_{not\ completed} = .37$, $t(156) = 1.09$, $p > .05$).

Psychological Well-being

The Psychological Well-Being (PWB; Ryff & Keyes, 1995) scale was used to measure well-being. This scale consists of 42 items measuring six subscales of adult psychological health: self-acceptance, positive relations with others, autonomy, environmental mastery, purpose in life, and personal growth. Participants rated their agreement to items such as "Most people see me as loving and affectionate" (positive relations with others) on a Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*). The six subscales have been shown to be highly intercorrelated, and responses are typically averaged to generate a score of overall well-being (Ryff & Keyes, 1995). Participants completed this scale both at the life narrative interview and the 3-year follow up. PWB showed good reliability at both times (PWB: $\alpha_{Time1} = .93$, $\alpha_{Time2} = .94$). Both the cross-sectional and longitudinal measurements were incorporated, as demonstrating that language and redemption are or are not overlapping in both cases provides the strongest test of their relationship.

Life Narrative Interview

During the baseline assessment, participants were given an in-person life-story interview (e.g., McAdams, 1985) with a university faculty member or a trained graduate student. From the full life story, low-point scenes were chosen for the present analyses. These stories have a greater chance of containing redemption, as they often deal directly with moments of adversity (Cox & McAdams, 2014). Past studies have demonstrated that redemption coding of responses to a single question in the life narrative interview is predictive of a variety of well-being measures. In the low-point portion of the interview, the participant is given the instructions:

> Thinking back over your entire life, please identify a scene that stands out as a low point, if not the lowest point in your life story. This does not have to be the lowest point in your life. Even though this event is unpleasant, I [the interviewer] would appreciate your providing as much detail as you can about it. What happened in the event, where and when, who was involved, and what were you thinking or feeling? Also, please say a word or two about why you think this particular moment was so bad and what the scene may say about you or your life.

Interviewers could prompt participants if any of these questions were not answered in the initial telling of the story. Interviews were digitally recoded and sent to a professional transcription service, Voss Transcriptions. Coding occurred using the text documents generated by this service.

Redemption Coding

Traditionally, life narratives are coded for redemption using a coding scheme that indicates, for each episode in the life narrative, whether redemption is present or absent; the total number of episodes containing redemption is then used as the individual's score on redemption. Given that this study only used low-point stories, a traditional dichotomous scheme would have lost precision and increased statistical error. To create a better estimate of redemption, we used a different coding scheme.

The second author and a trained undergraduate research assistant coded the lowpoint stories for redemption; a redemption sequence was defined by movement from intensely negative content to positive content. Thus, for a narrative to be coded as containing redemption, the following elements needed to be present: 1) description of a negative psychological, interpersonal, or emotional state; 2) description of positive elements or states of affairs that flowed from or were intrinsically connected to the negative phenomena; and 3) the final state of affairs, or end of the story, included the positive elements, i.e., the story did not end in a negative state. Low points that did not contain these three elements were coded as a 0 for redemption. Low points that contained these three elements but in which the positive content was limited, circumscribed, or did not fully address the negative content, were coded as a 0.5 for redemption. Finally, if the low point contained the three above elements and the positive content was extensive, elaborative, or fully addressed all aspects of the negative aspects of the episode, the story was coded as a 1 for redemption. Inter-rater reliability was adequate (intraclass correlation coefficient [ICC] = 0.83, 95% confidence interval [CI] [0.77, 0.87]; calculated using the psych package in R [Revelle, 2015]). Redemption scores spanned the full range of scores ($M = 0.60$, $SD = .77$), with 90 participants (57%) receiving a score of 0, 40 participants receiving a score of 0.5 (25%), and 28 (18%) receiving a score of 1.

Analyses

Word use in the low-point scenes was analyzed with the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2001) program. The LIWC program uses a set of predefined categories that contain specific words and word stems (for documentation, see Tausczik & Pennebaker, 2010). For this study, categories not related to specific words—including *words per sentence* and *punctuation*—were excluded (with the exception of *total word count*), as were categories that measure nonwords, such as *fillers* and *assents*. The LIWC program output provides percentages of a text that falls into a given category. For example, a score of 6.54 on *personal pronouns* indicates that 6.54% of the words in the text are in the personal pronoun category.

Given the exploratory nature of this study, all remaining word categories were included in the correlations between LIWC variables and redemption scores. However, not all categories could be used in the principal components analyses because LIWC output is hierarchical in nature. In other words, some categories are composed of other categories in the output. For example, the category *negative emotion* is the sum of the categories *anger*, *anxiety*, and *sad*. This hierarchy will systematically affect any data-reduction procedures, and thus higher order categories were not included in the principal components analysis. Removed categories were *function words*, *pronouns*, *personal pronouns*, *verbs*, *social*, *affect*, *negative emotion*, *cognitive/mechanical*, *perception*, *biological*, and *relativity*. Categories were also removed if they were sparsely used, here defined as having been used less than 1% of the time by fewer than 10% of the participants. This was done to reduce the inclusion of skewed variables. The following categories were sparsely used: *swearing*, *friend*, *anxiety*, *inhibition*, *sexual*, *ingestion*, *money*, and *religion*. In the following component analyses, a total of 43 linguistic categories were used. Those categories are listed in Table 1.

A parallel analysis was used to determine the appropriate number of components to extract from the principal components analysis. Parallel analysis is a simulation method which computes the eigenvalues for many sets of random, normally distributed data, and compares those eigenvalues with those observed in the actual data (Horn, 1965). Eigenvalues greater than 95% of the simulated eigenvalues are retained. Using the psych package (Revelle, 2015) for R (R Core Team, 2014), eigenvalues from the observed data set were plotted against the 95th-percentile eigenvalues of randomly resampled data and randomly generated data.

Once the appropriate number of components was determined, principal components analysis was performed to determine the content of those components. We chose principal components over factor analysis in order to capture as much of the variability in word use as possible (Jolliffe, 2005; Revelle, 2015). We extracted components using an oblimin rotation (for loadings, see Table 1). Correlations between components ranged from .00 to .18 ($M_r = .08$, $SD_r = .06$).

STUDY 1 RESULTS

How Do Categories of Language Relate to Redemption?

To determine whether word use is associated with redemption, each LIWC category was correlated with redemption. Few LIWC word categories were significantly associated with the redemption coding. As individuals increased in their redemption score, they constructed longer stories; that is, redemption was positively associated with *word count* ($r = .17$, 95% CI [.02, .32]). As predicted, both *positive emotion* words ($r = .21$, CI [.06, .36]) and *negative emotion* words ($r = –.19$, CI[–.34, –.04]) were significantly correlated with redemption. These particular findings indicate that as individuals devote more relative space to positive emotion words and less relative space to negative emotion words, they are more likely to be coded as having a redemptive sequence. Additionally, redemption was positively associated with three LIWC categories: *hear* ($r = .20$, CI [.05, .35]), *achievement* ($r = .16$, CI [.00, .16]), and *religion* ($r = .19$, CI [.04, .34]). For all correlations between redemption and word use, see Supplementary Tables S1.

After adjusting for multiple comparisons with a Holm adjustment (Holm, 1979), no correlations between redemption coding and linguistic categories were significant. Overall, only a few LIWC categories were associated with redemption, with moderate, at best, magnitudes, suggesting that redemption coding cannot be easily captured by any single word category.

Are There Patterns of Word Use in Life Narratives?

Because word use is a complex behavior, we next examined linear combinations of word-use categories, derived from principle components analysis. Principle components analysis was run on select LIWC variables to examine patterns of word use, as these patterns might better capture the combination of features that define redemption sequences. Parallel analysis demonstrated that seven components had eigenvalues greater than 95% of eigenvalues from simulated data, and the scree plot showed a drop at seven components. Thus, both methods converge on the finding that seven components is most appropriate for the sample (see Figure S1 for a depiction of the parallel analysis).

Table 1 Principal Components Solution for Study 1 Low-Point Stories

	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7
insight	0.73						
I	0.73						
negation	0.58						
number	−0.47						
article	−0.43		−0.41				
feel	0.42						
home		0.69					
impersonalpronoun		−0.64					
quantitative		−0.52					
tentativity		−0.44					
motion		0.43					
hear		0.42					
conjunction			0.66				
space			−0.59				
prepositions			−0.55				
family			0.50				
death			0.45				
you				0.77			
present tense				0.67			
past tense				−0.57			
inclusion				−0.43			
anger					0.57		
adverb					−0.57		
sad					−0.52		
exclusion					−0.50		
they					0.45		
word count					0.41		
future tense						0.84	
discrepancy						0.75	
health							0.68
body							0.66
she/he							0.45
work							−0.43
leisure							−0.42
Eigenvalue	3.40	3.25	2.86	2.78	2.75	2.50	2.22

Note. Pattern matrices of the rotated solutions are presented, as these facilitate interpretation. Loadings can be interpreted as the standardized regression weights predicting components from linguistic variables. Only standardized loadings greater than or equal to .40 are shown.

Table 1 includes the composition of the seven components and the loadings of variables on those components. Generally, we found patterns of word use that may capture both content of low-point stories and the ways in which these stories were told. For the sake of simplicity, we describe only some components here: The first component was characterized by use of *insight*-related words, *first-person pronouns* ("I," "me," "mine") and *negations*. The fifth component was characterized by increased use of *anger* and decreased use of *adverbs*, and *sadness*. The sixth component was characterized of *future-tense verbs* and *discrepancy* terms (e.g., "besides," "ought"). Consistent with previous use of data-reduction techniques in automated linguistic analysis (Chung & Pennebaker, 2008; Pennebaker & King, 1999), these results suggest that there are patterns of word use that may be extracted from low-point stories and analyzed in the study of life narratives.

Does the Structure of Word Categories Relate to Redemption Scores?

Given that the word categories can be grouped into components, we may find that these components are highly overlapping with redemption sequences. In other words, while a few word categories demonstrated tenuous associations with redemption coding, perhaps a linear combination of categories found in the principal components analysis is a better candidate for overlap with redemption coding. From a conceptual standpoint, Component 6 is the most similar to redemption in that both include use of linear time and describe contrasts. Specifically, redemption moves through at least a negative and to a positive event; Component 6 includes future-tense verbs and discrepancy words, which relate to desired versus undesired states.

To test the relationship of word-use components and redemption coding, component scores were calculated for each participant. Component scores were calculated by multiplying the standardized regression weights (presented in Table 1) with standardized scores on the linguistic variables. (See Supplementary tables S1 for correlations between components and the demographic covariates of gender and education.)

These component scores were compared to the redemption scores. None of the components were significantly associated with redemption (mean $r = .08$, ranging from .01 to .15; $p > .05$). Together with the adjusted correlations of LIWC variables, this suggests that word use in a low-point narrative is mostly independent of a redemptive story. In other words, redemption coding cannot be substituted with automated linguistic analysis, nor can automated linguistic analysis be substituted with coding for redemption.

Do Redemption Coding and Automated Linguistic Analysis Each Provide Unique Utility to the Study of Life Narratives?

Redemption coding and word-use components are largely independent of one another. If both redemption coding and automated linguistic analysis are individually useful in the prediction of important outcomes, then their joint application has the potential to enrich our understanding of life narratives. Both redemption coding and some word-use components predicted the outcome of psychological well-being. (See Supplementary tables S1 for these correlations.)

Given both redemption and some word-use components predicted well-being, using multiple regressions we sought to determine whether they were independent predictors. Redemption and the extracted language components were hierarchically entered into a multiple regression to predict psychological well-being at baseline (see Table 2). When controlling for age, gender, and education, redemption significantly predicted psychological well-being at the time of interview (*std b* = .27, CI [.13, .43]). Adding redemption to the covariate-only model (age, gender, and education) increased R^2 from .06 to .14, ($p < .05$), indicating that redemption coding more than doubled the explained variance of well-being. Each linguistic component was subsequently added to the model and removed, to determine the unique predictive ability of each. Only Component 6 was a significant predictor of well-being (*std b* = .26, CI [.11, .40]) when redemption was in the model. When this component was added, the R^2 value increased from .14 to .20, ($p < .05$), indicating that Component 6 adds as much predictive validity to the model as redemption coding. Redemption and Component 6 remained significant when all language variables were entered into the model indicating that they are unique predictors of psychological well-being. The R^2 of the full model was .24 ($p < .05$).

For a more stringent test, a similar series of models was run for the prediction of psychological well-being 3 years after baseline. After controlling for demographics, redemption significantly predicted psychological well-being at the 3-year follow-up (*std b* = .24, CI [.08, .41]). As before, each linguistic component was added to and removed from this model. Both Component 5 (*std b* = .19, CI [.02, .36]) and Component 6 (*std b* = .23, CI [.07, .38]) were significant predictors of psychological well-being at the 3-year follow up, and in both, redemption remained a significant predictor. Including these variables increased R^2 from .09 to .12 ($p < .05$) and .14 ($p < .05$), respectively. Components 5 and 6 and redemption coding were each significant predictors when all linguistic components were entered into the model, and the final R^2 value was .19 ($p < .05$). In sum, redemption coding cannot be captured by linguistic analyses, and both methods of analyses provide predictive validity.

STUDY 1 DISCUSSION

Overall, human narrative coding of redemption appears to be largely unrelated to word categories in low-point narratives. However, previous studies examining the relationship between word use and individual differences also find small average correlations (e.g., $r = .08$; Yarkoni, 2010). In fact, many of the associations in Study 1 are comparable to the upper limit of correlations found between LIWC categories and personality traits (e.g., $r = .24$; Hirsh & Peterson, 2009). While the significance of the correlation tests did not hold up to multiple corrections, we may have limited power to detect true but modest effects.

These effect sizes in this study demonstrate that any one linguistic category explains very little of the human narrative coding of redemption. Additionally, the combination of many linguistic categories, derived through the use of exploratory principal components analysis, was unrelated to the human narrative coding of redemption. Moreover, human narrative coding of redemption and linguistic components were independent predictors of psychological well-being 3 years later, suggesting that these measures reflect different, but important, psychological constructs. These results suggest that automated linguistic analysis and human narrative coding are minimally related, at least with respect to narrative redemption.

STUDY 2

The previous study tested the relationship between two methods of studying life narratives—automated linguistic analysis and trained human narrative coding of redemption—and found that these methods capture relatively dissimilar information. However, word use is largely influenced by contextual factors (Chung & Pennebaker, 2007), and the context of the Study 1 sample is geographically and developmentally particular. The independence of human narrative coding for redemption and word use, and larger structures of word use, may be specific to this late-midlife adult sample. We employed a second sample to attempt to replicate the structure of word use and association with redemption coding from Study 1. This study did not contain a measure of psychological well-being, so we are unable to replicate these results.

Table 2 Multiple Regression Predicting Psychological Well-being, Concurrently and Longitudinally in Study 1

Age	Gender	Education	Redemption	C1	C2	C3	C4	C5	C6	C7	R^2
Psychological well-being at interview											
.11(.08)	**.23(.08)**	.05(.08)									.06
.14(.08)	**.22(.08)**	.06(.08)	**.28(.08)**								.14
.13(.08)	**.25(.08)**	.06(.08)	**.27(.08)**	–.07(.08)							.14
.13(.08)	**.22(.08)**	.07(.08)	**.27(.08)**		.02(.08)						.14
.14(.08)	**.22(.08)**	.06(.08)	**.28(.08)**			.01(.08)					.14
.13(.08)	**.22(.08)**	.06(.08)	**.27(.08)**				.04(.08)				.14
.14(.08)	**.23(.08)**	.08(.08)	**.26(.08)**					.11(.08)			.15
.15(.07)	**.18(.07)**	.03(.07)	**.27(.07)**						**.26(.07)**		.20
.14(.08)	.21(.08)	.10(.08)	**.30(.08)**							.15(.08)	.16
.14(.08)	**.22(.09)**	.06(.08)	**.27(.08)**	–.09(.08)	–.03(.08)	.01(.08)	.05(.08)	.11(.08)	**.26(.07)**	.15(.08)	.24
Psychological well-being at 3-year follow-up											
–.16(.30)	**.18(.08)**	–.01(.09)									.03
–.05(.29)	**.16(.08)**	.00(.09)	**.24(.08)**								.09
–.07(.30)	.14(.09)	.00(.09)	**.24(.08)**	.07(.09)							.09
–.05(.30)	.16(.08)	.00(.09)	**.24(.08)**		.00(.09)						.09
–.04(.29)	.15(.08)	.01(.09)	**.25(.08)**			.11(.09)					.10
–.06(.30)	**.17(.08)**	.00(.09)	**.24(.08)**				.03(.09)				.09
–.11(.29)	**.17(.08)**	.01(.08)	**.21(.08)**					**.19(.09)**			.12
–.01(.29)	.13(.08)	–.02(.08)	**.24(.08)**						**.23(.08)**		.14
–.04(.29)	.16(.08)	.02(.09)	**.26(.08)**							.09(.09)	.10
–.07(.29)	.11(.09)	.01(.09)	**.23(.08)**	.04(.09)	–.03(.09)	.09(.09)	.02(.09)	**.19(.09)**	**.20(.08)**	.09(.08)	.19

Note. All coefficients are standardized. Bold coefficients are significant at $p < .05$.

STUDY 2 METHODS

Participants

Participants ($N = 242$) were recruited through the SAPA-Project.org website (Condon, 2014). Of the 242 participants whose low-point narratives were included in these analyses, 222 (61.7% female) had complete background data. These participants were younger than the sample from Study 1 and had a wider spread of ages ($M_{age} = 29.00$, $SD_{age} = 13.03$). These participants were also slightly less diverse: 46% of the sample was White, 2% Black, 5% Asian/Asian American, 6% Hispanic/Latino, and 7% Other/Mixed Race; 34% of the sample did not specify their ethnicity. Approximately 31% of the sample had a college degree (with another 14% completing some college and 39% currently attending a college or university), making the sample fairly well educated, similar to Study 1.

Life Narrative Interview

Participants were given a set of instructions that asked them to describe either a high point in their life or a low point in their life, similar in form to Study 1. The instructions given were randomly assigned, with 66% of participants receiving a low-point story. The prompt for this story was written to be as close as possible to the prompt given in the interviewer format:

> Thinking back over your entire life, please identify a scene that stands out as a low point, if not the low point in your life story. *This does not have to be the lowest point in your life, but merely a very bad experience of some kind*. In the following scene, include these details: What happened in the event? Where and when did the event take place? Who was involved? What were you thinking and feeling? Why do you think this particular moment was so bad? What does this scene say about who you are as a person?

For these analyses, only stories collected through the low-point instructions were used, so as to more closely match the stories analyzed in the previous sample. There were a total of 327 low-point stories. However, some included only a sentence or two. To ensure the highest quality data, stories with fewer than 100 words were excluded from the analyses, leading to the 242 low-point stories used in these analyses.

Analyses

Two trained research assistants coded these stories for redemption. Each coder rated the story as containing no redemption or redemption, and we averaged their responses to create a total redemption score for each participant. Inter-rater reliability was adequate (ICC = 0.77, 95%CI [.75, .79]). One-hundred twenty-two (122) participants (50%) received a final redemption score of 0, 36 participants (15%) received a score of 0.5, and 84 (35%) received a score of 1.

Language was assessed with the LIWC program. Both individual categories and components were used in the analyses.

Principal components analysis was used to identify the linguistic structure—an identical component analysis plan (seven components, oblimin rotation) to Study 1. The same variables that were removed from Study 1 (e.g., categories which contained other categories and *swearing*, *friend*, *anxiety*, *inhibition*, *sexual*, *ingestion*, *money*, and *religion*) were excluded from the principal components analysis, and seven components were extracted. It should be noted that visual inspection of the scree plot suggested that three, five, and seven components could be appropriate for these data. This initial test implies that the structure found in Study 1 is unlikely to fit the data in Study 2 well.

STUDY 2 RESULTS

How Do Categories of Language Relate to Redemption?

Redemption scores were first correlated with word categories. Some correlations found in Study 1 were replicated: individuals high in redemption wrote more words ($r = .14$, 95% CI [.02, .26]), used more *positive emotion* words ($r = .14$, [.02, .27]) and described *achievement* ($r = .21$, [.09, .33]). Additionally, individuals high in redemption used more *numbers* ($r = .13$, [.00, .25]) and words related to *health* ($r = .15$, [.03, .27]) and *work* ($r = .21$, [.08, .32]); these individuals used fewer *impersonal pronouns* ($r = –.14$, [–.26, –.01]), *negations* ($r = –.13$, [–.25, .00]) and words related to *anger* ($r = –.15$, [–.27, –.02]) and *insight* ($r = –.14$, [–.26, –.01]). After adjusting for multiple comparisons, none of these correlations remained significant. Overall, these results in conjunction with the findings from Study 1 suggest that redemption may be weakly associated with number of words used, positive emotion words, and achievement words, but is generally unassociated with word-use categories overall.

What Is the Structure of Word Categories in Life Narratives?

Linguistic components were extracted under the same conditions as in the previous study (i.e., using principal components, seven components, oblimin rotation); results are shown in Table 3. Correlations between components ranged from .00 to .18 ($M_r = .06$, $SD_r = .05$). Again, we describe only select components here: The second component included high loadings of *exclusions*, *tentative*, and *negations*. The third component was characterized by *achievement*- and *work*-related words and decreased use of words relating to *hearing*. The sixth component was characterized by *inclusion* and *conjunctions*. Components will be referred to as Component 1b, etc., to distinguish from components in the previous study. As in Study 1, we were able to extract components of word use, which may capture general content themes or writing styles.

The components found in Study 2 are wholly unrelated to the components found in Study 1. For example, Component 4 found in Study 1 included use of *you*-words and *present tense*; neither of these word categories loaded strongly onto any component in Study 2. Thematically, these components were not replicated, e.g., a health component was found in Study 1, but an achievement component was found in Study 2. Overall, the structure of word use found in Study 1 did not replicate in a broader population, indicating that contextual factors likely have an impact on word use patterns.

Does the Structure of Word Categories Relate to Redemption Scores?

In Study 1, no language components were associated with redemption. By contrast, one language component in Study 2 was significantly associated with redemption scores: Component 3b ($r = .21$, [.08, .33]). This association is likely driven by the relationship between redemption scores and achievement-word use, which may be due to the relationship between redemption and agency (Adler et al., 2015; Baerger & McAdams, 1999; McAdams et al., 2001).

STUDY 2 DISCUSSION

Human narrative coding of redemption was again largely unrelated to word categories and related to only one linguistic component. It is worth noting that human narrative coding of redemption was positively associated with *achievement* and *positive emotion* words, replicating associations found in Study 1. These findings suggest that word use is, at best, marginally related to whether a life narrative is considered redemptive by human coders. In addition, the results further support the finding that human narrative coding of redemption cannot be captured by an aggregate of word-use categories, as a single linguistic component was modestly related to human narrative coding of redemption. The linguistic components showed little similarity between the two samples, which may be attributable to sample differences, method differences between the two studies, or a true null result.

GENERAL DISCUSSION

The two studies directly compared the two dominant forms of narrative analysis, automated linguistic analysis and human narrative coding, by examining the overlap between word use and redemption in low-point narratives. These are the first studies to directly compare these methods when applied to the life narrative and thus serve as test of whether or not these methods are redundant. Together, these studies converge on the finding that human narrative coding of redemption cannot be captured by automated linguistic analysis, as the empirical relationship between human narrative coding of redemption and word categories was modest at best. Both human narrative coding of redemption and patterns of word use independently predicted future levels of well-being, suggesting that both methods have unique utility in the study of life narratives. In sum, our findings

Table 3 Principal Components Solution for Study 2 Low Points

	Component 1b	Component 2b	Component 3b	Component 4b	Component 5b	Component 6b	Component 7b
prepositions	0.69						
auxiliary verb	−0.68						
space	0.67						
past tense	−0.57						
exclusion		0.80					
tentativity		0.76					
negatation		0.61					
present tense		0.42					
achievement			0.76				
work			0.72				
hear			−0.44				
home			−0.43				
family			−0.42				
positive emotion			0.40				
insight				0.58			
feel				0.51			
wordcount				−0.50			
she/he				0.49			
causality				−0.41			
I					−0.62		
we					0.56		
quantitative					0.53		
discrepancy					−0.42		
inclusion						0.78	
conjunction						0.70	
humans							-0.58
sad							0.51
time							0.50
Eigenvalue	2.83	2.80	2.56	2.35	2.18	2.06	2.06

Note. Only loadings greater than or equal to .40 shown.

indicate that redemption coding describes emergent qualities of life narratives that are unable to be captured merely through word use, while automated linguistic analysis describes the elements from which life narratives are constructed but do not capture broader narrative arcs. Thus, these two methods should be considered complementary approaches to the study of life narratives and should be used simultaneously to best understand the important psychological qualities that are reflected in a life story.

Despite the general nonconvergence between the two methods, human narrative coding of redemption was modestly associated with some word categories. Consistent with expectations, there is an association between redemption and *positive emotion* words, likely a result of redemption sequences being partly defined by the description of positive events or outcomes. Additionally, human narrative coding of redemption was associated with achievement words in both studies. While the association with *achievement* was not specifically hypothesized, a greater focus on achievement suggests these individuals are describing more agency-based redemption (Adler et al., 2015; Baerger & McAdams, 1999; McAdams et al., 2001). Agency-based redemption occurs when a participant reports enhanced self-efficacy, strength, confidence of self-understanding as a result of the change from positive to negative. Achievement words as measured by LIWC may be an example of participant-reported enhanced self-efficacy. This finding highlights the power of the simultaneous use of these two methods of narrative analysis, as automated linguistic analysis identified empirical support for a relationship between redemption and agency, a theorized component of redemption (McAdams et al., 2001). Interestingly, the relationships found between human narrative coding of redemption and word use were similar in magnitude to found relationships between LIWC output and other individual differences (e.g., personality traits; Hirsh & Peterson, 2009; Kern et al., 2014; Yarkoni, 2010). Consistent associations of this magnitude from study to study suggest that word use as measured by LIWC may, at best, correlate with commonly studied individual differences at modest levels. On the one hand, word use may not be strongly associated with these individual differences, suggesting that speaking and writing patterns reflect new, unique aspects of one's personality that need to be further examined. On the other, it is possible that alternative methods and techniques of measuring word use (e.g., aggregation across multiple contexts) are needed to uncover larger associations.

The methodological import of this study is noteworthy, as these findings suggest that simultaneous use of automated linguistic analysis and human narrative coding should be encouraged. There are advantages to employing both human narrative coding and automated linguistic analysis as each method offers a different set of benefits. While human coders allow researchers to simultaneously examine the thematic qualities of life narratives, automated linguistic analysis efficiently captures details of life narratives that are too subtle to be consciously processed (Chung & Pennebaker, 2007). For example, human coders are able to detect differences in narrative arcs, while automated linguistic analysis can pick up on the differences in use of specific articles (e.g., "the," "an") that are unconsciously processed and consequently difficult to incorporate into the rating of an emergent property. Additionally, human narrative coding allows for developing conceptually complex coding procedures to test specific theories (such as the examination of growth-oriented motivational themes in the prediction of well-being; Lilgendahl & McAdams, 2011), while automated linguistic analysis is amenable to exploratory statistical procedures to uncover patterns of word use (such as patterns of words which represent individual differences in writing styles; Chung & Pennebaker, 2008; Pennebaker & King, 1999). We feel it is important to use both of these methods because life stories are rich with information, with the typical study currently only able to assess a fraction of that information (Adler et al., 2015). Beyond the additional information gained by using these multiple methods, an additional advantage of life stories is apparent when compared to the typical individual difference study. A multi-method goal or trait study needs to collect data from multiple sources (e.g., self- and informant reports; Jackson et al., 2015), whereas an advantage of collecting life narratives is the ability to use multiple methods with the same stimuli (i.e., the narrative itself).

This study offers an example of the benefits of utilizing both human narrative coding and automated linguistic analysis in the study of life narratives. Our results show that word use and human narrative coding of redemption independently predict psychological well-being 3 years after the life story interview, indicating that word use and redemption influence well-being through different mechanisms. Word use is thought to reflect broad individual differences in thinking style (Chung & Pennebaker, 2008; Pennebaker & King, 1999), such that use of specific words is indicative of framing the world in a certain way. In our own study, individuals who used more future tense experienced greater well-being. Use of future-tense words suggests thinking about and planning for the future, which leads to better well-being (MacLeod, Coates & Hetherton, 2008). Thus, our study suggests that when narrating life story low points, a future orientation is beneficial. Redemption, on the other hand, is a result of meaning making, or integrating negative events into a coherent life story (McLean & Pratt, 2006), which in some cases leads to better adjustment and overall well-being (Greenhoot & McLean, 2013; McAdams et al., 2001; McAdams & Guo, 2015). Together, these findings suggest that future studies consider the independent roles of thinking styles and meaning-making processes in the construction of a life story.

In this way, linguistic analysis can be a vital tool for narrative psychologists. The components that emerged in the present studies represent qualities that can distinguish between narratives in a given sample. These components point to features that may be specific to a population or to a type of story. Such features allow narrative researchers to discover previously unconsidered themes, generate new hypotheses for exploration (as proposed above), and more fully understand the narratives they have collected.

Despite including two studies that utilized two different methods, the current investigation had a number of limitations. First, while the LIWC program is perhaps one of the most widely used and most user-friendly methods of automated linguistic analysis, it does not encompass the entire field of automated linguistic analysis. Other forms of analysis can examine life stories at much finer grained levels (e.g., the differential language approach; Eichstaedt et al., 2015) or can incorporate the ways in which words and phrases are used together (e.g., Riley et al., 2005). Such methods of automated linguistic analysis may be able to provide meaningful information both further removed from the emergent level of human narrative coding and, perhaps, more similar. One major limitation of all these methods is the inability to capture sequential elements within a text. Second, this study was limited in its scope of life story elements (i.e., low points), themes of interest (i.e., redemption) and outcomes (i.e., well-being). Future research should cast a broader net to understand whether the relationship between human narrative coding of themes and word use depends on the story told or the construct of interest, and to determine whether there are outcomes for which one method of analysis is a better predictor. Third, this study considered only a limited number of analyses when assessing the relationship between word use and redemption. Given that our aim was to assess the degree to which two common methods overlap, we chose to use these methods and analyze their results in the ways another life narrative researcher would likely use them. However, there are a number of other techniques that may consider more complex relationships between words and redemption.

The current study advances the field of life narrative and personality research by starting to illuminate the relationship between the two dominant methods for investigating narratives. We conclude that human narrative coding of redemption is not simply the sum of its parts, as it is not strongly associated with word use; instead, redemption is an emergent property of life stories that must be measured and interpreted as an independent construct. Additionally, this study shows that automated linguistic analysis can add to the understanding of life stories, and thus future research should incorporate this method of analysis alongside the more common method of human narrative coding. Research into method and related theory is necessary for the field of life narrative and personality to mature. As such, it is vital that life story researchers embrace both approaches: automated linguistic analysis and human narrative coding.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was in part funded by the Foley Foundation's support of the Foley Center for the Study of Lives, Northwestern University.

References

Adler, J.M. (2012). Living into the story: Agency and coherence in a longitudinal study of narrative identity development and mental health over the course of psychotherapy. *Journal of Personality and Social Psychology*, **102**, 367–389.

Adler, J.M., Lodi-Smith, J., Philippe, F.L., & Houle, I. (2015). The incremental validity of narrative identity in predicting well-being: A review of the field and recommendations for the future. *Personality and Social Psychology Review*, doi: 10.1177/1088868315585068.

Adler, J.M., & Poulin, M.J. (2009). The political is personal: Narrating 9/11 and psychological well-being. *Journal of Personality*, **77**, 903–932.

Baerger, D.R., & McAdams, D.P. (1999). Life story coherence and its relation to psychological well-being. *Narrative Inquiry*, **9**, 69–96.

Benish-Weisman, M., Wu, L.M., Weinberger-Litman, S.L., Redd, W.H., Duhamel, K.N., & Rini, C. (2014). Healing stories: Narrative characteristics in cancer survivorship narratives and psychological health among hematopoietic stem cell transplant survivors. *Palliative and Supportive Care*, **12**, 261–267.

Chung, C. & Pennebaker, J.W. (2007). The psychological function 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.

Condon, D.M. (2014, July 28). *The SAPA Project*. Retrieved from http://www.sapa-project.org

Cox, K., & McAdams, D.P. (2014). Meaning making during high and low point life story episodes predicts emotion regulation two years later: How the past informs the future. *Journal of Research in Personality*, **50**, 66–70.

Dunlop, W.L., & Tracy, J.L. (2013a). Sobering stories: Narratives of self-redemption predict behavioral change and improved health among recovered alcoholics. *Journal of Personality and Social Psychology*, **104**, 576–590.

Dunlop, W.L., & Tracy, J.L. (2013b). The autobiography of addiction: Autobiographical reasoning and psychological adjustment in abstinent alcoholics. *Memory*, **21**, 64–78.

Eichstaedt, J.C., Schwartz, H.A., Kern, M.L., Park, G., Labarthe, D. R., Merchant, R.M., et al. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science*, doi: 10.1177/0956797614557867.

Graybeal, A., Sexton, J.D., & Pennebaker, J.W. (2002). The role of story-making in disclosure writing: The psychometrics of narrative. *Psychology and Health*, **17**, 571–581.

Greenhoot, A.F., & McLean, K.C. (2013). Introduction to this Special Issue Meaning in personal memories: Is more always better? *Memory*, **21**, 2–9.

Hammack, P.L. (2008). Narrative and the cultural psychology of identity. *Personality and Social Psychology Review*, **12**, 222–247.

Hirsh, J.B., & Peterson, J.B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality*, **43**, 524–527.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.

Horn, J. (1965). A rationale and test for the number of components in component analysis. *Psychometrika*, **30**, 179–185.

Jackson, J.J., Connolly, J.J., Garrison, S.M., Leveille, M.M., & Connolly, S.L. (2015). Your friends know how long you will live: A 75-year study of peer-rated personality traits. *Psychological Science*, **3636,** 335–340. doi:10.1177/0956797614561800

Jolliffe, I. (2005). *Principal component analysis*. Hoboken, NJ: John Wiley & Sons, Ltd.

Josselson, R., & Lieblich, A. (1993). *The narrative study of lives*. Thousand Oaks, CA: Sage.

Kern, M.L., Eichstaedt, J.C., Schwartz, H.A., Dziurzynski, L., Ungar, L.H., Stillwell, D.J., et al. (2014). The online social self: An open vocabulary approach to personality. *Assessment*, **21**, 158–169.

King, L.A. (2001). The health benefits of writing about life goals. *Personality and Social Psychology Bulletin*, **27**, 798–807.

Lilgendahl, J.P., & McAdams, D.P. (2011). Constructing stories of self-growth: How individual differences in patterns of autobiographical reasoning relate to well-being in midlife. *Journal of Personality*, **79**, 391–428.

Lodi-Smith, J., Geise, A.C., Roberts, B.W., & Robins, R.W. (2009). Narrating personality change. *Journal of Personality and Social Psychology*, **96**, 679–689.

MacLeod, A.K., Coates, E., & Hetherton, J. (2008). Increasing well-being through teaching goal-setting and planning skills: Results of a brief intervention. *Journal of Happiness Studies*, **9**, 185–196.

Manczak, E.M., Zapata-Gietl, C., & McAdams, D.P. (2014). Regulatory focus in the life story: Prevention and promotion as expressed in three layers of personality. *Journal of Personality and Social Psychology*, **106**, 169–181.

McAdams, D.P. (1985). *Power, intimacy, and the life story*. Homewood, IL: Dorsey.

McAdams, D.P. (1995). What do we know when we know a person? *Journal of Personality*, **63**, 363–396.

McAdams, D.P. (2006). The problem of narrative coherence. *Journal of Constructivist Psychology*, **19**, 109–125.

McAdams, D.P., Albaugh, M., Farber, E., Daniels, J., Logan, R.L., & Olson, B. (2008). Family metaphors and moral intuitions: How conservatives and liberals narrate their lives. *Journal of Personality and Social Psychology*, **95**, 978–990.

McAdams, D.P., Anyidoho, N.A., Brown, C., Huang, Y.T., Kaplan, B., & Machado, M.A. (2004). Traits and stories: Links between dispositional and narrative features of personality. *Journal of Personality*, **72**, 761–784.

McAdams, D.P., Diamond, A., de St. Aubin, E., & Mansfield, E. (1997). Stories of commitment: The psychological construction of generative lives. *Journal of Personality and Social Psychology*, **72**, 678–694.

McAdams, D.P., & Guo, J. (2015). Narrating the generative life. *Psychological Science*, **26**, 475–483.

McAdams, D.P., Reynolds, J., Lewis, M., Patten, A.H., & Bowman, P.J. (2001). When bad things turn good and good things turn bad: Sequences of redemption and contamination in life narrative and their relation to psychosocial adaptation in midlife adults and in students. *Personality and Social Psychology Bulletin*, **27**, 474–485.

McAdams, D.P., & Pals, J.L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, **61**, 204–217.

McLean, K.C., & Breen, A.V. (2009). Processes and content of narrative identity development in adolescence: Gender and well-being. *Developmental Psychology*, **45**, 702–710.

McLean, K.C., Pasupathi, M., & Pals, J.L. (2007). Selves creating stories creating selves: A process model of self-development. *Personality and Social Psychology Review*, **11**, 262–278.

McLean, K.C., & Pratt, M.W. (2006). Life's little (and big) lessons: Identity statuses and meaning-making in the turning point narratives of emerging adults. *Developmental Psychology*, **42**, 714–722.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Mahway, NJ: Lawrence Erlbaum Associates.

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., Mayne, T.J., & Francis, M.E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology*, **72**, 863–871.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Revelle, W. (2015). *psych: Procedures for Personality and Psychological Research*. Evanston, IL: Northwestern University. Retrieved from http://CRAN.R-project.org/package=psych

Riley, K.P., Snowdon, D.A., Desrosiers, M.F., & Markesbery, W.R. (2005). Early life linguistic ability, late life cognitive function, and neuropathology: Findings from the Nun Study. *Neurobiology of Aging*, **26**, 341–347.

Robbins, M.L., Mehl, M.R., Smith, H.L., & Weihs, K.L. (2013). Linguistic indicators of patient, couple, and family adjustment following breast cancer. *Psycho-Oncology*, **22**, 1501–1508.

Rude, S., Gortner, E.M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, **18**, 1121–1133.

Ryff, C.D., & Keyes, C.L.M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, **69**, 719–727.

Singer, J.A. (2004). Narrative identity and meaning making across the adult lifespan: An introduction. *Journal of Personality*, **72**, 437–460.

Tausczik, Y.R., & Pennebaker, J.W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, **29**, 24–54.

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, **44**, 363–373.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1 Parallel analysis scree plots.

Supplementary Tables S1 Contains additional comparisons to address relationships between variables not discussed in the manuscript. Included are the correlations between individual LIWC categories and psychological well-being, and the correlations between the linguistic components and demographics and well-being.