

# Patterns of student writing in a critical thinking course: A quantitative analysis

David W. Carroll\*

*Department of Human Behavior and Diversity, Old Main 324, University of Wisconsin-Superior,  
Superior, WI 54880-4500, USA*

---

## Abstract

Two studies assessed student writing in a critical thinking course. The first study examined initial and final versions of a term paper. Linguistic inquiry and word count (LIWC) analyses of essays revealed significant differences in linguistic variables and some differences in cognitive variables. The second study examined student essays at the beginning and end of the semester. Taken together, the studies indicate that students express themselves with less tentativeness at the end of the semester, and are consistent with the suggestion that students move from relativism to commitment during the course of a semester. In addition, the results indicate that the LIWC may detect differences between different writing genres.

© 2008 Elsevier Inc. All rights reserved.

*Keywords:* Writing assessment; Linguistic inquiry and word count (LIWC); Critical thinking; Personal epistemology; Intellectual growth; Psychology

---

## 1. Introduction

Instructors ask students to undertake writing assignments for a variety of reasons. In some instances, the purpose is to increase students' mastery of a particular body of knowledge. On other occasions, the goal is primarily to foster the students' writing skills. In still other contexts, the purpose is to foster intellectual growth, such as an increase in critical thinking skills. Some writing assignments may involve all of these goals.

In this paper, I will examine the influence of a course in critical thinking on college students' cognitive and linguistic growth. I will first consider different theories of intellectual growth in college students and identify both similarities and differences in these treatments. I will then

---

\* Tel.: +1 715 394 8323; fax: +1 715 394 8493.

*E-mail address:* [dcarroll@uwsuper.edu](mailto:dcarroll@uwsuper.edu).

discuss two studies that describe patterns of language use in a critical thinking course. Finally, I will consider to what extent these patterns fit the aforementioned theories of intellectual growth.

## 2. Literature review

Theoretical treatment of intellectual growth in late adolescence and adulthood have variously been described as research in epistemological beliefs, reflective judgment, ways of knowing, epistemological reflection, and epistemological resources (Hofer, 2001). It appears that the most common umbrella term for such theories is personal epistemology (Hofer & Pintrich, 2002). Theories of personal epistemology attempt to explain how adolescents and young adults think about knowledge—that is, how they assess knowledge claims as well as how individuals use their conceptions of knowledge to develop an understanding of the world. In addition, many theories of personal epistemology are developmental in nature and seek to explain how changes in conceptions of knowledge may be related to experiences in and out of the classroom. We will look at several theories: Perry (1970), Belenky, Clinchy, Goldberger, and Tarule (1986), King and Kitchener (1994), and Baxter Magolda (1992).

An early and influential treatment of intellectual growth in college students was the work of Perry (1970). Perry conducted open-ended interviews with students at Harvard in the 1960s and used these observations to develop a stage theory of intellectual growth during the college years. Each stage was described in terms of its dominant epistemological position. Students do not invariably move through the sequence in a linear order, as there was assumed to be pausing and backward movement as well as forward movement. Moreover, there are assumed to be individual differences in the speed with which students move through the stages.

Perry identified nine positions, but there were three main stages. Perry's first stage was dualism. Dualists see the world in terms of right/wrong, black/white, and good/bad. Dualists are passive learners who depend upon authorities to hand down the truth. Later, when exposed to different points of view, students become aware of the multiplicities of perspectives and more relativistic in their point of view. At this stage, they understand that different individuals have different points of view, but do not see a principled basis for evaluating these viewpoints. Later, after they have been challenged by their teachers' insistence on evidence, they arrive at the commitment within relativism stage, where they come to adopt a consistent point of view but with appropriate justification as well as awareness of alternative viewpoints.

Belenky et al. (1986) expanded on Perry's formulation by focusing on women's epistemological views. Belenky et al. noted that although Perry interviewed both men and women, only men's interviews were used in illustrating and validating his theoretical scheme. Perry then examined whether women's interviews also fit with the stages derived from men's interviews, and concluded that they did. Belenky et al. agreed, but also noted that there may be other routes that women take in addition to those they share with men.

Belenky and colleagues interviewed 135 women from colleges such as Bard College, La Guardia Community College, and other settings (such as social services clients engaged in improving parenting skills). The interviews included sections dealing with self-concept, moral judgment, relationships and educational experiences as well as a section that has come to be called "ways of knowing" (Clinchy, 2002).

Belenky et al.'s interviews suggested five categories (not necessarily stages): silence, received knowledge, subjective knowledge, procedural knowledge, and constructed knowledge. The most conspicuous distinction from Perry's scheme was the silence category, a position in which women feel disconnected from knowledge and power. Women may resort to silence when they do not feel

safe in expressing beliefs that challenge authorities in the classroom (Schrader, 2004). Although Belenky et al. initially characterized the silence category as rare, Clinchy (2002) has more recently questioned this point.

Women in the received knowledge category emphasize the value of receiving and reproducing knowledge. These women have little confidence in their own voice and instead associate learning with the ability to receive, retain, and return the words of various authorities. This category is similar to Perry's dualism position (Clinchy, 2002) and based on the assumption that truth is absolute and unambiguous.

The subjective knowledge category is associated with distrust for logic, analysis, and abstraction, and an increasing interest in developing one's own voice. Instead of relying upon external authorities, the subjective knower believes that truth lies within. Truth springs from the heart or gut. First-hand experience is a valuable source of knowledge.

Women in the procedural knowledge category no longer believe that knowledge consists of facts to be acquired. Rather, they believe that knowledge is a process and requires work (Clinchy, 2002). Moreover, although no single answer may be correct, all interpretations are not equally valid. The procedural knowledge category has elements of two different forms of knowing, separate knowing and connected knowing. Connected knowing is based in one's capacity for empathy whereas separate knowing is based on the separation from feelings and emotions in the cause of objectivity.

The distinction between connected and separate knowing is an aspect of Belenky et al.'s work that has generated considerable discussion as well as a fair amount of misunderstanding. Separate knowing is not the problem—it is a form of detached, scientific thinking commonly observed in the sciences. However, Clinchy (2002) suggests that some writers have confused connected knowing with subjectivism, a view that emphasizes effortless intuition. In contrast to the reflexive, spontaneous subjectivism, connected knowing is deliberate and effortful.

Finally, constructed knowledge integrates these different strands. It involves "weaving together the strands of rational and emotive thought and of integrating objective and subjective knowing" (Belenky et al., p. 134). Constructive thinkers show a high tolerance for internal contradiction and ambiguity, and an emphasis on dialogue and balance. Once the knower assumes the general relativity of knowledge, they can construct and reconstruct different frames of reference and evaluate particular claims with regard to various standards or criteria.

Another approach to personal epistemology is represented by the work of King and Kitchener (1994, 2002, 2004). Following Dewey (1933), King and Kitchener view reflective judgment as a process by which a person brings closure to an uncertain situation. In particular, King, Kitchener and colleagues have examined how late adolescents and adults make judgments about controversial and ill-structured issues, such as whether the benefits of inoculating health care workers against smallpox will outweigh the risks or whether affirmative action is an effective tool for promoting genuine access to higher education. These and similar issues are problems about which reasonable people may disagree.

The reflective judgment model (King & Kitchener, 2002, 2004; Kitchener & Fischer, 1990) includes seven stages of reflective judgment, beginning with stages in which knowledge is assumed to exist absolutely and concretely and progressing to stages in which knowledge is considered to be the outcome of a process of reasonable inquiry in which solutions to ill-structured problems are constructed. At the final stage, the adequacy of a solution is evaluated by considering the available evidence and is reevaluated when new perspectives, evidence, or tools of inquiry become available.

The reflective judgment model has been the subject of extensive research. Most of the research has been based on the Reflective Judgment Interview (RJI), a one-hour semi-structured discussion

of four ill-structured problems. King and Kitchener (1994) review the responses of over 1,700 participants who have completed the RJI. Using longitudinal data, King and Kitchener (2002) conclude that the reflective judgment stages form a developmental sequence. King, Kitchener, and Wood (1994) found that the RJI scores of virtually all participants (92%) increased over a 10-year period. Moreover, most of the participants followed the stages of the model, with few examples of stage skipping or regressions.

Cross-sectional studies have examined the influence of educational level and gender on reflective judgment. King et al. (1994) found that high school students consistently showed pre-reflective thinking whereas traditional-age college students tend to display more quasi-reflective answers. Graduate students consistently earn the highest RJI scores. Thus, it appears that educational experience plays a significant role in RJI performance. In contrast, gender differences on RJI are less reliable, and those that occur may be attributed to differences in educational attainment.

It should also be mentioned that the RJI is only one approach to assessing reflective judgment. Kitchener, Lynch, Fischer, and Wood (1993) compared performance on the RJI with a variant called the Prototypic Reflective Judgment Interview (PRJI), an interview in which participants are given additional contextual support in the form of probe questions. The purpose of the contextual support was to assist individuals in reaching their optimal level of reflective judgment. The results indicated that participants reached a higher stage of reflective judgment when given additional contextual support.

Both Perry and Belenky et al. began their investigations without specifically focusing their studies in personal epistemology. Like King and Kitchener, Baxter Magolda (1992, 2002) began a longitudinal study of the epistemological development of both women and men college students. Baxter Magolda identified four major ways of knowing, each of which is composed of two different and equally complex reasoning patterns. The ways of knowing were found to evolve from absolute, in which knowledge is viewed as certain, through transitional, in which some knowledge is regarded as certain to independent knowing, in which knowledge is uncertain because everyone has their own beliefs, and finally culminating in a contextual way of knowing that depends upon the evaluation of contextually significant evidence to assess various knowledge claims.

Although these theories diverge in various ways, there is agreement that there are different ways of knowing that emerge between adolescence and young adulthood (see West, 2004). Both Perry and Kitchener and colleagues emphasize stages of development more than Belenky et al., who identify categories. Even for those theories that stipulate stage-like development, there are assumed to be considerable individual variations in the rate of development and sometimes the order or sequencing of the stages. The epistemological positions that adolescents and young adults assume may vary with educational opportunity, context, and support. Still, there is broad agreement that adolescent thinking often emphasizes dualism, absolute truth, and received knowledge whereas young adults often view knowledge in a reflective, constructive, and committed manner. Young adults are increasingly capable of taking positions that are more individualistic, integrate different perspectives, and respect evidence.

This paper examines the intellectual changes in student writing during a critical thinking class. The overriding goal of the course is to encourage students to view phenomena from multiple perspectives. The specific approach taken is to examine “weird things” such as astrology, parapsychology, communicating with the dead, ghosts, and the like. Students often enter the class with strong beliefs on many of these topics. Both class presentations and the primary textbook for the class (Schick & Vaughn, 2005) encourage students to examine their assumptions and consider alternative explanations of these events.

More specifically, the class encourages students to consider the limitations of their own experiences. Class presentations on visual illusions, memory lapses, and judgment errors demonstrate that it is possible to strongly believe that one has witnessed an unusual event and still be wrong. By systematically examining alternative hypotheses, students are encouraged to consider that there are multiple possible explanations for many events and that their own particular interpretation is not necessarily the correct one. The goal of the course is not necessarily to encourage students to change their belief systems (although it appears that some do). Rather, the goal is to stimulate students to examine the evidence for their beliefs. Put another way, the goal is to encourage students to bring their level of confidence in their beliefs in line with the degree of evidence. Because the evidence is unclear in many instances, some students may leave the class with more questions than answers.

One example is interpreting near-death experiences (NDEs). Individuals who survive heart attacks sometimes report experiences that include entering a dark tunnel, seeing a source of light, and experiencing peacefulness. Those experiencing NDEs may also have an out-of-body experience, in which they look down at their own body from the ceiling of an emergency room. Most interestingly, patients can sometimes report events and conversations that occurred when they were unresponsive. One example is reporting, in great detail, the appearance of a nurse that entered the emergency room after the patient became unconscious. A common interpretation of these experiences is that individuals are experiencing a glimpse of the afterlife before being pulled back into this life. Many students in the critical thinking class accept this interpretation, apparently without question, at the beginning of the class.

Class discussion of NDEs is designed to examine other points of view. The sense of peacefulness may be based on the release of endorphins, which are chemicals in the body that act similarly to morphine. The experience of dark and light might be related to the impairment of brain functioning if the oxygen supply to the brain has been affected. The ability of patients to report conversations that occurred when they were unresponsive might be explained by natural events. For example, the patient might have encountered the nurse in a previous visit to the hospital.

Once alternative hypotheses for a phenomenon have been established, students are encouraged to compare their adequacy by using a variety of criteria. Essentially, students are trained in scientific reasoning, and to view the “best” hypothesis as the one that can explain the most phenomena with the fewest number of assumptions.

There were five exams in the course, and six homework assignments, all designed to assess student mastery of the concepts in the primary or secondary (Epstein & Kernberger, 2006) textbooks. However, the centerpiece of the course was a term paper in which students applied the principles of critical thinking to a topic of their interest. The project was completed in stages. During the first week of the semester, students identified a topic of interest and discussed what previous beliefs or experience, if any, they have had with the topic. Later, they supplied an annotated reference list and completed a rough draft of their paper. The final paper was due at the end of the semester. They received graded feedback on all of these assignments.

Studies of personal epistemology have tended to use a combination of interviews and questionnaires. Perry (1970) invited participants based on their scores on the Checklist of Educational Values to participate in interviews regarding their educational experiences. Aspects of Perry’s original questionnaire are present in a widely used current measures of personal epistemology, such as Schommer’s Epistemological Beliefs Questionnaire (Hofer, 2002; Schommer, 1990) and the Epistemic Belief Inventory (Schraw, Bendixen, & Dunkle, 2002). Both Belenky et al. and Baxter Magolda have used refined interview protocols that have drawn upon Perry’s questions.

King and Kitchener's (1994) approach is slightly different, as they have interviewed individuals about ill-structured problems; Kuhn (1991) has used a similar approach.

In this manuscript, I approach the development of knowledge beliefs through the quantitative analysis of student essays. Although past research has often relied on interviews and questionnaires to assess intellectual development, both methods have their limitations. Interview and questionnaire responses may be influenced by biases introduced by the presence of an interviewer and the desire of participants to please interviewers or testers. In contrast, text analysis assesses a writer's words without influencing the words. Moreover, it is often useful to combine multiple methodologies to better assess the validity of individual methodological approaches.

In recent years, a number of methods of quantitative analysis of text have developed and have been applied to a variety of research problems (Mehl, 2006). In this study, I examine whether the Linguistic Inquiry and Word Count (LIWC) software developed by Pennebaker and associates (Pennebaker, Francis, & Booth, 2001; Pennebaker, Mehl, & Niederhoffer, 2003) may detect changes in student essays from the beginning to the end of the semester. LIWC receives text files as input and produces measures of 74 linguistic and psychological variables. Most measures are the percentages of various word types (e.g., articles, emotion words) in the text. In addition, researchers may develop customized categories of personal interest.

A number of studies have explored the relationships between LIWC measures and variables such as age (Pennebaker & Stone, 2003), personality (Pennebaker & King, 1999), and depression (Stirman & Pennebaker, 2001). For example, Stirman and Pennebaker (2001) examined poems written by writers who committed suicide compared to a group matched in age and other variables. Suicidal poets used more first person singular pronouns than nonsuicidal poets, which is consistent with theories that posit that self-preoccupation is a factor in suicide. Rude, Gortner, and Pennebaker (2004) found similar results in a sample of college students.

Although Pennebaker and colleagues have examined the usefulness of the LIWC in a wide variety of contexts, the software has not been applied often to academic writing. Pennebaker and King (1999) examined the psychometric properties of the LIWC by analyzing samples of academic work of undergraduate students. This study found that the LIWC displayed high levels of reliability across different writing samples. In addition, Pennebaker and King used academic writing to determine the factor structure of the LIWC. However, I have not found any LIWC studies that examined changes in LIWC as a function of academic experience.

I am most interested in the LIWC variables that pertain to linguistic and cognitive words. The linguistic and cognitive variables used in this study are shown in Tables 1 and 2. Although the distinction between these two categories may appear to be somewhat arbitrary, considerable LIWC research demonstrates that the two categories of variables are related to different authors, genres, and writing conditions (Pennebaker et al., 2003). Words per sentence and big words tend to be higher in more academic or technical writing than in personal writing. Words in dictionary, a measure of the percentage of words in a text that are captured by the LIWC analysis, tends to be lower in more technical writing; that is, technical writing introduces terms not in the standard dictionary more often than personal writing. Pronouns are less common in academic than personal writing.

In contrast, cognitive variables are associated with writing about emotional events. For example, people who write about emotional upheavals have improved psychological and physical health (Pennebaker & Francis, 1996). This pattern is particularly true for individuals whose writing includes high levels of cognitive words (*think, know, consider*) as well as positive emotion words (*cheerful, fun, pleasant*) with a moderate level of negative emotion words (*angry, anxious, sad*). Thus, cognitive words appear to be a sensitive index of the extent to which a person is "processing" a particular event, especially an emotional event.

Table 1  
Linguistic variables used in Studies 1 and 2

Variable	Definitions and examples
Word count	Number of words in text
Words per sentence	Number of words per sentence
Questions	Percent of sentences with questions
Unique words	Percent of words
Words in dictionary	Percent of text words captured by LIWC dictionary
Big words	Percent of words longer than six letters
Pronouns	<i>I, my, me</i>
Negatives	<i>No, never, not</i>
Articles	<i>A, an, the</i>
Prepositions	<i>On, to, from</i>
Numbers	<i>One, thirty, million</i>

Table 2  
Cognitive variables used in Studies 1 and 2

Variable	Examples
Causation	<i>Because, effect, hence</i>
Insight	<i>Think, know, consider</i>
Discrepancy	<i>Should, would, could</i>
Inhibition	<i>Block, constrain</i>
Tentativeness	<i>Maybe, perhaps, guess</i>
Certainty	<i>Always, never</i>

This paper reports two studies that examine student intellectual development using the method of quantitative text analysis. The first study examines student writing on weird topics at the beginning and the end of the semester. The second compares early and late student essays on a topic not discussed in class.

### 3. Study 1

For the first study, I examined 17 variables, 11 linguistic and six cognitive. I chose the linguistic variables because previous research studies found that percentage of various word categories (e.g., pronouns, articles) is sensitive to differences between writers, writing tasks, or both. I chose cognitive variables (e.g., words related to certainty, tentativeness, and insight) because I thought that a course in critical thinking might influence the use of such words. I looked at student essays at the beginning and at the end of the semester.

In general, I expected students' final papers to reflect the main themes of the course: the search for objective evidence, the examination of alternative explanations, and an appreciation of scientific reasoning. I also anticipated that their essays would show intellectual growth similar to that proposed by various personal epistemology theorists (Baxter Magolda, 1992; Belenky et al., 1986; King & Kitchener, 1994; Perry, 1970). However, even if we grant that students, as these theorists have supposed, generally progress from dualism to relativism to commitment during their college years, it is not clear what stage of growth students are typically in at the beginning of the semester. One could conjecture that many students enter the course with a predominantly

dualistic perspective, and move toward relativism and the appreciation of multiple perspectives during the course of the semester. However, it is also possible that students begin the semester with a relativistic framework and move toward commitment.

Two specific hypotheses were tested. The first hypothesis was that linguistic and cognitive variables would show significant change from the beginning to the end of the semester. The second hypothesis was that the initial measure of the variables showing significant change would correlate with student performance at the end of the semester.

### 3.1. Method

#### 3.1.1. Participants

Forty-two students enrolled in a sophomore-level course in critical thinking in Fall 2003 and 2004 participated in the study. The course was cross-listed in Philosophy and Psychology, and roughly half of the students were in each section. Many of the students were first-term freshmen.

#### 3.1.2. Materials and procedure

The primary text for the course was a book that encouraged students to critically examine a number of “weird” things, such as parapsychology, alternative medicine, and creationism (Schick & Vaughn, 2005). The term paper for the course required students to apply principles of critical thinking to a “weird” topic of their choice. Students completed the paper in a series of drafts, but only the first and last versions were included in this study. The initial essay required students to identify a topic of interest and explain why it interested them. The final essay was a term paper (10–15 pages) that applied principles of critical thinking to the topic of choice.

Students submitted their papers electronically as Word attachments. I converted the papers into text files and eliminated names, dates, and lists of references, and then submitted the papers to the LIWC analysis. Tables 1 and 2 show examples of the cognitive and linguistic variables used in the analysis. To evaluate the second hypothesis, student performance was defined as the number of points that students earned on the five exams in the course.

### 3.2. Results

An 11 (Variable)  $\times$  2 (Time) repeated measures ANOVA on the linguistic variables revealed a significant effect for Variable,  $F(10, 410) = 455.73$ ,  $p < .001$ , partial  $\eta^2 = .92$ , Time,  $F(1, 41) = 335.52$ ,  $p < .001$ , partial  $\eta^2 = .89$ , and the Variable by Time interaction,  $F(10, 410) = 343.13$ ,  $p < .001$ , partial  $\eta^2 = .89$ . Paired  $t$  tests revealed that seven of the variables showed a significant difference between pre- and post-test scores (see Table 3). Initial essays showed shorter word counts, fewer unique words, fewer dictionary words, more words greater than six letters (“big words”), fewer pronouns, more articles, and more numbers ( $p < .001$ ). There was also a trend toward an increase in words per sentence ( $p < .10$ ). There were no differences in the use of questions, negatives, or prepositions.

A 6 (Variable)  $\times$  2 (Time) repeated measures ANOVA on the cognitive variables revealed a significant effect for Variable,  $F(5, 205) = 123.41$ ,  $p < .001$ , partial  $\eta^2 = .75$ , Time,  $F(1, 41) = 8.76$ ,  $p < .005$ , partial  $\eta^2 = .18$ , and the Variable by Time interaction,  $F(5, 205) = 5.87$ ,  $p < .001$ , partial  $\eta^2 = .13$ . Final essays had fewer insight ( $p < .001$ ) and discrepancy ( $p < .05$ ) words, and more words related to inhibition ( $p < .05$ ). There were trends ( $p < .10$ ) toward more words related to cause and fewer words related to tentativeness (see Table 4).

Table 3  
Linguistic variables for initial and final essays in Study 1

Variable	Initial <sup>a</sup>		Final		<i>t</i>
	<i>M</i>	S.D.	<i>M</i>	S.D.	
Word count	403.54	172.62	3389.90	1080.50	18.51**
Words per sentence	21.29	4.68	22.71	4.43	1.92*
Questions	3.40	7.37	3.68	3.14	.24
Unique words	51.32	5.82	28.22	3.66	24.29**
Words in dictionary	72.72	6.30	67.41	4.69	6.72**
Big words	20.12	4.49	23.84	4.12	7.15**
Pronouns	9.82	3.76	5.68	2.35	9.52**
Negatives	1.30	.68	1.42	.37	1.19
Articles	7.14	1.95	8.30	1.18	4.51**
Prepositions	13.81	2.04	13.49	1.08	.89
Numbers	1.22	.79	2.17	.76	5.62**

<sup>a</sup>  $n = 42$ .

\*  $p < .10$ .

\*\*  $p < .001$ .

Two multiple regression analyses were conducted to predict student exam points. One analysis included the eleven linguistic variables as predictors and the second analysis included the six cognitive variables. The regression equation for the linguistic variables was significant,  $R^2 = .42$ , adjusted  $R^2 = .24$ ,  $F(11, 33) = 2.21$ ,  $p < .05$ . The regression equation for the cognitive variables was not significant,  $R^2 = .15$ , adjusted  $R^2 = .02$ ,  $F(6, 38) = 1.14$ ,  $p > .10$ . A multiple regression analysis conducted with all 17 measures as predictors approached significance,  $R^2 = .52$ , adjusted  $R^2 = .21$ ,  $F(17, 27) = 1.73$ ,  $p < .10$ .

Correlational analyses explored the relationship between all 17 variables in the initial essays and the number of exam points students achieved. There were significant correlations between initial essay word count ( $r = .40$ ), dictionary words ( $r = -.38$ ), big words ( $r = .38$ ) and exam points. Partial correlations indicated that initial word count was correlated with exam points when big words and dictionary words were partialled out,  $r(41) = .35$ ,  $p < .05$ . Neither big words nor percent of dictionary words predicted exam points when word count was partialled out, although dictionary words approached significance,  $r(41) = -.27$ ,  $p < .10$ .

Table 4  
Cognitive variables for initial and final essays in Study 1

Variable	Initial <sup>a</sup>		Final		<i>t</i>
	<i>M</i>	S.D.	<i>M</i>	S.D.	
Causation	1.47	.81	1.72	.60	1.86*
Insight	3.42	1.26	2.77	.70	3.65***
Discrepancy	1.85	1.03	1.47	.51	2.49**
Inhibition	.24	.31	.36	.13	2.31**
Tentativeness	2.75	1.29	2.37	.70	1.85*
Certainty	1.16	.58	1.07	.33	1.25

<sup>a</sup>  $n = 42$ .

\*  $p < .10$ .

\*\*  $p < .05$ .

\*\*\*  $p < .001$ .

### 3.3. Discussion

The present study explored two hypotheses. The first hypothesis was that linguistic and cognitive variables would show significant change from the beginning to the end of the term. This hypothesis was clearly supported for linguistic variables and moderately for cognitive variables. Although the word count difference was required for students, other linguistic variables appear to indicate real change. At the end of the semester, students were more likely to use words not in the Pennebaker dictionary, big words, articles, and numbers, and were less likely to use pronouns. Cognitive variables showed fewer significant changes throughout the semester than linguistic variables. Students were less likely to use words related to insight, discrepancy, and tentativeness, and more likely to express inhibition and causal thinking.

Some of these changes may be related to the critical thinking course. The tendency to use non-dictionary words may reflect increased attention to technical terms (such as endorphins) that were related to their topics. Moreover, the increase in numbers may be due to greater sensitivity to issues of evidence. Students were required to find and evaluate evidence for the different hypotheses they considered. Accordingly, final papers more often included results, including quantitative results, from various research studies.

Similarly, the decline in personal pronouns may reflect course content. As noted earlier, the class emphasizes increased skepticism regarding one's personal experience, and thus a decline in personal pronouns was expected. Moreover, the shift was particularly clear for the first-person singular (5.1% of all words in initial essays, 1.1% in final papers). It is likely that the decline in pronouns and corresponding increase in articles is related to students moving from using personal experience to research evidence as sources of support for their beliefs. Moreover, there are some style manuals, such as American Psychological Association style, that de-emphasize personal voice and emphasize objectivity in writing (Madigan, Johnson, & Linton, 1995). To the extent that students are exposed to these emphases, it is likely that their use of pronouns will decline over the course of their college career, at least when writing scientific papers.

The most interesting finding in the cognitive variables was that students were less likely to use words related to tentativeness (such as *maybe*, *perhaps*, and *guess*) in their final essays than in their initial essays. In addition, students were less likely to use insight words (such as *think*, *know*, and *consider*) in their final essays. The significant decline in insight words may be related to the observation that cognitive words increase with cognitive or emotional uncertainty (Pennebaker & Francis, 1996). Pennebaker and Francis found that cognitive words appear more often in essays of persons who have undergone an emotional upheaval, such as a health crisis or the loss of a loved one (Pennebaker & Lay, 2002). Students likely had greater anxiety or uncertainty regarding their topics at the beginning of the semester, and thus prefaced more comments with *I think* than in their final papers. In addition, my students made a number of comments during class discussions that expressed an awareness of multiplicity and relativism. In contrast, their final papers seemed to display a level of commitment, at least on the topic of choice. These cognitive changes are consistent with theoretical descriptions of the intellectual growth that students often undergo during their college experience.

There is, however, a significant problem in interpreting these results because the two essays differ not only in the time of writing but also in their nature and requirements. The initial assignment required no research or sources and encouraged students to share their experiences related to the topic. The final assignment was a traditional term paper. Thus, the two assignments differed in genre. There is a need to tease apart the genre and temporal differences by exploring changes over time in essays with similar requirements.

The second hypothesis was that the initial measure of significant variables would be correlated with student exam performance at the end of the semester. This hypothesis was partially supported. The initial measure of 3 of the 10 variables that showed significant change over the semester correlated with student exam performance. Of the three, initial word count and number of dictionary words were most predictive of exam performance. Students who wrote longer initial essays and used more uncommon words performed better on exams. Although limited, these results suggest that initial essays can detect those students who are better prepared to be successful in a critical thinking class. Students who write longer initial essays in introductory emails have been found to be more conscientious than other students and perform better in classes (Carroll & Biscay, 2001). Thus, initial student writing samples may be indicative of final class performance in some instances.

## 4. Study 2

A second study was conducted to examine changes in student writing without the influence of genre. For this study, students were asked to write a one-page essay on what makes people happy. This question is similar to the ill-structured problems employed by other researchers (King & Kitchener, 1994; Kuhn, 1991). Since the topic of happiness was not discussed in the class, this essay permits us to examine whether the changes observed in Study 1 generalize from a familiar topic to a less familiar one.

### 4.1. Method

#### 4.1.1. Participants

Nineteen students enrolled in a sophomore-level course in critical thinking in Fall 2006 participated in the study.

#### 4.1.2. Materials and procedure

In the first week of class, students were given a sheet with a single question, “What makes people happy?” and were asked to write a one-page essay. Students wrote a second essay near the end of the semester without seeing their earlier one. Students wrote the essays anonymously. They put an alphanumeric code on their papers to facilitate pairing of initial and final essays.

### 4.2. Results

An 11 (Variable)  $\times$  2 (Time) repeated measures ANOVA on the linguistic variables revealed a significant effect for Variable,  $F(10, 180) = 598.88$ ,  $p < .001$ , partial  $\eta^2 = .97$ , Time,  $F(1, 18) = 10.76$ ,  $p < .01$ , partial  $\eta^2 = .37$ , and the Variable by Time interaction,  $F(10, 180) = 12.36$ ,  $p < .001$ , partial  $\eta^2 = .41$ . Paired  $t$  tests revealed that two of the variables showed a significant difference between pre- and post-test scores. Initial essays showed larger word counts and fewer unique words ( $p < .01$ ) (see Table 5).

A 6 (Variable)  $\times$  2 (Time) repeated measures ANOVA on the cognitive variables revealed a significant effect for Variable,  $F(5, 90) = 34.13$ ,  $p < .001$ , partial  $\eta^2 = .66$  and the Variable by Time interaction,  $F(5, 90) = 4.02$ ,  $p < .002$ , partial  $\eta^2 = .18$ . The Time effect approached significance,  $F(1, 18) = 3.42$ ,  $p < .10$ , partial  $\eta^2 = .16$ . Paired  $t$  tests revealed that final essays had fewer words related to tentativeness ( $p < .001$ ) and a trend for more words related to certainty ( $p < .10$ ) (see

Table 5  
Linguistic variables for initial and final essays in Study 2

Variable	Initial <sup>a</sup>		Final		<i>t</i>
	<i>M</i>	S.D.	<i>M</i>	S.D.	
Word count	148.47	30.02	119.95	27.78	3.67**
Words per sentence	16.48	4.00	15.62	2.43	.92
Questions	.20	.48	.08	.26	1.43
Unique words	58.65	4.89	63.94	9.24	3.01**
Words in dictionary	79.70	2.70	78.40	5.03	1.02
Big words	17.54	3.62	17.89	4.18	.29
Pronouns	8.27	2.75	8.12	2.89	.19
Negatives	1.17	1.15	.87	1.04	1.22
Articles	5.55	1.81	5.68	2.70	.18
Prepositions	11.87	2.67	12.78	2.18	1.37
Numbers	.84	1.00	.62	.76	1.26

<sup>a</sup> *n* = 19.

\*\* *p* < .01.

Table 6  
Cognitive variables for initial and final essays in Study 2

Variable	Initial <sup>a</sup>		Final		<i>t</i>
	<i>M</i>	S.D.	<i>M</i>	S.D.	
Causation	1.27	1.32	1.21	.91	.17
Insight	2.04	.97	2.38	2.08	.63
Discrepancy	1.91	1.34	2.01	1.77	.17
Inhibition	.03	.14	.11	.34	.90
Tentativeness	5.29	2.46	2.92	2.57	2.92**
Certainty	.87	.67	1.33	1.02	1.76 <sup>c</sup>

<sup>a</sup> *n* = 19.

\* *p* < .10.

\*\* *p* < .01.

Table 6). Tentative words declined from 5.3% of all words to 2.9%, whereas certainty words increased from .8% to 1.3%.

#### 4.3. Discussion

There were two main findings in this study. First, the differences between initial essays and final essays were much less pronounced than in Study 1. Although there were eleven differences (seven linguistic and four cognitive) between initial and final essays found in Study 1, there were only four (two linguistic and two cognitive) in Study 2. This finding suggests that most of the differences found in Study 1 were more likely attributable to genre than to cognitive or linguistic growth.

Second, the decline in tentativeness between initial and final essays, although significant in Study 1, was much larger in Study 2. Moreover, there was a trend for an increase in words related to certainty in Study 2, whereas this variable produced no difference in Study 1. It thus appears that the tendency for students near the end of the semester to write in a less tentative manner reflects true cognitive growth. By using hedges and other forms of tentative speech less often, and words related to certainty somewhat more often, students seem to be identifying and clarifying

their own voice. It appears that students at the end of the class use language that is more consistent with a reflective, constructive, and committed view of knowledge claims.

## 5. General discussion

This paper examined student writing in a sophomore-level course in critical thinking with the intent of examining changes in student thinking as reflected in their language. The two studies taken together suggest the need to closely examine both genre and temporal differences in writing. Most of the differences in writing found in Study 1 (in particular, the tendencies to use bigger words, fewer pronouns, more articles, and more dictionary words) appear to be attributable to differences in the requirements of the two assignments. Although students wrote about the same topic in the two assignments (i.e., their topic of choice), they approached the first assignment as a personal essay and the latter one as an academic term paper, and the aforementioned linguistic variables reflected that.

In contrast, some of the cognitive differences found in Study 1—particularly, the decline of tentative writing—was even stronger in Study 2. This finding is striking in that the class never discussed the topic of happiness. It thus appears that the experience of learning and writing about critical thinking influences student writing, even about topics not specifically studied. Nor can the changes observed in Study 2 be attributed to the topic of the essays, since that variable was held constant. The most likely conclusion is that the cognitive changes reflect changes in students' epistemological views. The changes observed in this paper are consistent with the view that students begin the class with a relativistic or subjective view of knowledge and progress toward a constructive and committed view by the end of the semester.

Some examples from student essays may illustrate these shifts. Here is the first sentence of two students' essays at the beginning of the semester:

In life, there are many things that make people happy.

There can be a great deal of things that can and will make many individuals happy.

At the end of the semester, these students began their essays with the following sentences, respectively:

The thing that makes people happy is to satisfy needs.

People are happy when they have others around them who are happy.

These students' initial essays emphasized the diversity of ways individuals might answer this question. In contrast, their later essays began with a thesis that was subsequently justified. These essays appear to reflect the relativity and commitment positions, respectively. Although these students' final essays continued to reflect an awareness of the diversity of possible answers to this question, they also demonstrated an inclination to discover some common patterns among the diverse responses. This attention to patterns was conspicuously absent in the essays earlier in the semester.

The present studies are limited in a number of ways. The course consisted primarily of freshmen in their first semester of college. It is not clear whether similar changes would occur in classes other than critical thinking or whether we would see such changes from the beginning to the end of an upper-division course. The changes observed in Study 2 might be due to adolescent maturation, life experiences, and other factors beyond completing a course in critical thinking.

In addition, it is possible that students in Study 2 viewed the end-of-course essays as a means of evaluating their learning, and that this presupposition may have influenced their writing. That is, students may have been trying to impress their instructor with the extent to which they could express themselves with a more committed point of view.

These limitations suggest further studies. It would be interesting to examine changes in student writing in a senior-level course. Although most developmental theories allow for significant variation in the attainment of different positions, it might be expected that more seniors would have adopted the committed epistemological viewpoint at the beginning of the semester, and hence fewer cognitive changes over the course of the term would be observed. It would be also be useful to tease apart education from other developmental factors by examining pre- and post-test essays by young adults who are not enrolled in college.

The techniques used in this study have wide applicability. The LIWC is easy to learn to use, and can be applied to essays in a wide variety of disciplines. Moreover, it is easy to customize the LIWC variables for particular purposes. Text analysis systems such as LIWC, Coh-Metrix (Graesser, McNamara, Louwerse, & Cai, 2004) and the Landscape Model (van den Broek, Young, Tzeng, & Linderholm, 1999) do not replace qualitative analyses by instructors. However, they complement such evaluations by providing a fast, comprehensive analysis on a large number of linguistic variables (see Mehl, 2006).

In summary, the present studies found that quantitative analyses of student essays detect differences in types of writing assignments as well as changes in student writing throughout the semester. The results indicate that students are less tentative at the end of the semester than at the beginning. These studies suggest that quantitative analysis of student essays may be a useful complement to qualitative analysis.

## Acknowledgements

I would like to thank Jamie Pennebaker, Blaine Peden, and Todd Jackson for their comments on this manuscript.

## References

- Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey-Bass.
- Baxter Magolda, M. B. (2002). Epistemological reflection: The evolution of epistemological assumptions from age 18 to 30. In: B. Hofer & P. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 89–102). Mahway, NJ: Erlbaum.
- Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing: The development of self, voice, and mind*. New York: Basic Books.
- Carroll, D. W., & Biscay, J. (2001, May). *Electronic mail and class performance*. Poster session presented at the Council of Undergraduate Teachers of Psychology, Chicago.
- Clinchy, B. M. (2002). Revisiting women's ways of knowing. In: B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 63–87). Mahway, NJ: Erlbaum.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: Heath.
- Epstein, R. L., & Kernberger, C. (2006). *The pocket guide to critical thinking* (3rd ed.). Belmont, CA: Thomson.
- Graesser, A. C., McNamara, D. S., Louwerse, M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers*, 36, 193–202.
- Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Journal of Educational Psychology Review*, 13, 353–383.

- Hofer, B. K. (2002). Personal epistemology as a psychological and educational construct: An introduction. In: B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 3–14). Mahwah, NJ: Erlbaum.
- Hofer, B. K., & Pintrich, P. R. (Eds.). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Erlbaum.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment*. San Francisco: Jossey-Bass.
- King, P. M., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In: B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37–61). Mahwah, NJ: Erlbaum.
- King, P. M., & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, 39, 5–18.
- King, P. M., Kitchener, K. S., & Wood, P. K. (1994). Research on the reflective judgment model. In: P. M. King & K. S. Kitchener (Eds.), *Developing reflective judgment* (pp. 124–202). San Francisco: Jossey-Bass.
- Kitchener, K. S., & Fischer, K. W. (1990). A skill approach to the development of reflective thinking. In: D. Kuhn (Ed.), *Contributions to human development: Vol. 21. Developmental perspectives on teaching and learning thinking skills* (pp. 48–62). New York: S. Karger.
- Kitchener, K. S., Lynch, C. L., Fischer, K. W., & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. *Developmental Psychology*, 29, 893–906.
- Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.
- Madigan, R., Johnson, S., & Linton, P. (1995). The language of psychology: APA style as epistemology. *American Psychologist*, 50, 428–436.
- Mehl, M. R. (2006). Quantitative text analysis. In: M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 141–156). Washington, DC: American Psychological Association.
- Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion*, 10, 601–626.
- Pennebaker, J. W., Francis, M. E., & Booth, M. J. (2001). *Linguistic inquiry and word count*. Mahwah, NJ: Erlbaum.
- 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., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, ourselves. *Annual Review of Psychology*, 54, 547–577.
- Pennebaker, J. W., & Stone, L. D. (2003). Words of wisdom: Language use over the life span. *Journal of Personality and Social Psychology*, 85, 291–301.
- Perry, W. G., Jr. (1970). *Forms of ethical and intellectual development in the college years: A scheme*. New York: Holt, Rinehart, and Winston.
- Rude, S. S., Gortner, E.-M., & Pennebaker, J. W. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion*, 18, 1121–1333.
- Schick, T., Jr., & Vaughn, L. (2005). *How to think about weird things: Critical thinking for a new age* (4th ed.). Boston: McGraw Hill.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498–504.
- Schrader, D. E. (2004). Intellectual safety, moral atmosphere, and epistemology in college classrooms. *Journal of Adult Development*, 11, 87–101.
- Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). Development and validation of the epistemic belief inventory (EBI). In: B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 261–275). Mahwah, NJ: Erlbaum.
- Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63, 517–522.
- van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading. In: H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Mahwah, NJ: Erlbaum.
- West, E. J. (2004). Perry's legacy: Models of epistemological development. *Journal of Adult Development*, 11, 61–70.