

**IMPACT OF CATASTROPHE ON PIVOTAL NATIONAL LEADERS' VISION
STATEMENTS: CORRESPONDENCES AND DISCREPANCIES IN MORAL
REASONING, EXPLANATORY STYLE, AND RUMINATION**

A dissertation submitted
by

CARL RUSSELL OLIVER

to

FIELDING GRADUATE INSTITUTE

in partial fulfillment of
the requirements for the
degree of

DOCTOR OF PHILOSOPHY
in

HUMAN AND ORGANIZATIONAL SYSTEMS

This dissertation has been
accepted for the faculty of
Fielding Graduate Institute by:

Valerie Malhotra Bentz, Ph.D.
Chair

Judith Stevens-Long, Ph.D.
First Faculty Reader

Frank J. Barrett, Ph.D.
Second Faculty Reader

Kenneth R. Bartos, M.A.
Student Reader

Impact of Catastrophe on Pivotal National Leaders' Vision Statements: Correspondences and Discrepancies in Moral Reasoning, Explanatory Style, and Rumination

by
Carl Russell Oliver

Abstract

Vision statements articulated by 7 pivotal national leaders before and after a catastrophe were examined for changes in four latent characteristics: moral reasoning stage, moral reasoning orientation, explanatory style, and rumination. The 7 leaders are Abraham Lincoln, Woodrow Wilson, Franklin Roosevelt, John Kennedy, Golda Meir, Margaret Thatcher, and George W. Bush. Their respective catastrophes are the capture of Fort Sumter, Germany waging unrestricted submarine warfare, Japan's attack on Pearl Harbor, the Soviet Union placing offensive missile sites in Cuba, Egypt and Syria attacking Israel, Argentina attacking the Falkland Islands, and terrorists attacking the United States. Content structure analysis methods, inherently unobtrusive and non-reactive, were used for the multivariate study. An expert rater supported by validated computer software scored moral reasoning stage. The other three measures were scored by raters using two-stage protocols. In the face of catastrophe, these 7 leaders usually showed no significant change in moral reasoning stage, and neutral or positive explanatory style. They showed significant changes in rumination and in the caring and justice components of moral reasoning orientation. From an organizational systems perspective, an impact of human-caused catastrophe was adjustment of these four variables in pivotal national leaders' vision statements as if they are systems levers.

Key words: at-a-distance assessment, content analysis, explanatory style, hierarchical complexity, moral development, moral reasoning orientation, moral reasoning stage, organizational systems, rumination, shared leadership, structure analysis, unobtrusive measures.

Copyright by
CARL RUSSELL OLIVER
2004

Dare Website Edition, January 2007

Acknowledgments

Assistance is gratefully acknowledged. Michael Lamport Commons, Ph.D. graciously served on the dissertation committee as external examiner. T. L. Dawson, Ph.D., served as expert rater for hierarchical complexity. Research assistants from California Lutheran University, Thousand Oaks, California, who served as blind raters for the pilot study were undergraduates J. Card, M. Carlson, D. Daley, J. Rogers, and B. Zubiate; those who served as blind raters for the dissertation study were undergraduates J. Acquisti, T. D. Almer, J. Higgs, J. Jones, J. Magro, S. N. Mendivel, S. Miccio, and graduate students Y. V. Ribeiro, and P. G. Sitton. C. A. Oliver contributed to analysis of the pilot study and D. M. Oliver helped to enter raters' raw scores into computer spreadsheets.

Table of Contents

Chapter One:	Introduction.....	10
	Increasing Need for Catastrophe Leadership	11
	Vision Statement Criteria	12
	Problem Statement	12
	Key Definitions	13
	Pivotal Leaders	13
	Vision Statement.....	13
	Human-Caused Catastrophe	13
	Explanatory Style.....	13
	Measures of Explanatory Style	14
	Moral Reasoning Stage.....	14
	Measures of Moral Reasoning Stage	14
	Moral Reasoning Orientation	14
	Measures of Moral Reasoning Orientation.....	14
	Rumination	14
	Measures of Rumination.....	15
	Organizational Culture.....	15
	Prosocial Behavior.....	15
	Assumptions.....	15
	Research Question.....	15
Chapter Two:	Review of the Literature	16
	Introduction	16
	The Organizational Frame.....	16
	Leadership Research	17
	Origin and Trend.....	17
	Salient Research.....	18
	Application and Measure	21
	Social Psychology	23
	Origin and Trend: Organizational Culture.....	23
	Salient Research: Organizational Culture.....	23
	Application and Measure: Organizational Culture	23
	Origin and Trend: Prosocial Behavior	24
	Salient Research: Prosocial Behavior	25
	Origin and Trend: Explanatory Style.....	25
	Salient Research: Explanatory Style.....	27
	Application and Measures: Explanatory Style	27
	Developmental Psychology.....	28
	Origin and Trend: Moral Reasoning Stage.....	28
	Salient Research: Moral Reasoning Stage.....	29

	Application and Measures: Moral Reasoning Stage.....	30
	Origin, Trend and Salient Research: Moral Reasoning Orientation	32
	Application and Measures: Moral Reasoning Orientation	33
	Non-Reactive or Unobtrusive Measures	33
	Countering Reactivity	33
	Strengths and Weaknesses	33
	Content Analysis Facilitates Study of Catastrophe Leadership.....	34
	Summary	35
	Hypotheses	35
	Moral Reasoning Stage.....	35
	Moral Reasoning Orientation.....	36
	Explanatory Style and Rumination.....	36
Chapter Three:	Method.....	37
	Basic Research Design	37
	Selection of Leaders.....	37
	Identification of Vision Statements	37
	Unit of Coding	38
	Comparability and Levels of Analysis.....	38
	Extractions	39
	Reliability	39
	Experimental Texts	39
	Scoring.....	39
	Data Analysis	40
Chapter Four:	Findings.....	42
	Moral Reasoning Stage	42
	The Data Collected	42
	Reliability.....	44
	Statistically Significant Findings	45
	Anomalies	50
	Summary and Hypothesis Testing	50
	Moral Reasoning Orientation	51
	The Data Collected	51
	Reliability.....	53
	Significant Findings	53
	Anomalies	56
	Summary and Hypothesis Testing	58
	Explanatory Style	58
	The Data Collected	58

	Reliability	62
	Significant Findings.....	63
	Anomalies	66
	Summary and Hypothesis Testing	68
	Rumination	68
	The Data Collected	68
	Reliability	69
	Significant Findings.....	70
	Anomalies	73
	Summary and Hypothesis Testing	73
	Interactions	73
	Explanatory Style and Moral Reasoning Stage	74
	Moral Reasoning Orientation and Rumination.....	74
	Rumination and Explanatory Style.....	76
	Patterns Observed in the Data.....	76
	Factor Analysis	78
	Effect Size.....	79
	Summary	80
Chapter Five:	Discussion.....	81
	Summary	81
	Moral Reasoning Stage.....	84
	Moral Reasoning Orientation	84
	Explanatory Style.....	85
	Rumination	85
	Interactions	85
	Conclusions	86
	Methods	86
	Content.....	88
	Alternative Explanations	94
References	96

List of Tables and Figures

Table 1:	Pivotal Leaders and Catastrophes Selected for Study	10
Table 2:	Sample Behaviors for Moral Reasoning Stages 15 to 21	43
Table 3:	Number of Moral Reasoning Orientation Extractions for Each Vision Statement	52
Table 4:	Number of Explanatory Style Extractions for Each Vision Statement	59
Table 5:	Summary of Explanatory Style Ratings	60
Table 6:	One-Sample <i>t</i> -Test Results	63
Table 7:	Pre- and Post-Catastrophe Rumination Means Evaluated	70
Table 8:	<i>t</i> -test Paired-Comparison Correlation Matrix for Moral Stage, Moral Orientation, Explanatory Style, and Rumination	74
Table 9:	<i>t</i> -test Paired-Comparison Correlation Matrix for Rumination and Explanatory Style Components	76
Table 10:	Factor Loadings	79
Table 11:	Calculated Effect Sizes	80
Table 12:	Intercoder Reliability Measurements	87
Table 13:	Correlation of Explanatory Style Components to Each Other and to Composite Scores	93
Figure 1:	Stages of moral reasoning by vision, 1-3 before catastrophe and 4-6 afterward	42
Figure 2:	Frequency distribution of moral reasoning stages in the 41 scored visions	46
Figure 3:	Stages of moral reasoning by vision before and after a catastrophe	47
Figure 4:	Stages of moral reasoning for vision statements aggregated by period: pre- or post-catastrophe	49
Figure 5:	Frequency distribution of moral reasoning stages for vision statements aggregated by period: pre- or post-catastrophe	49
Figure 6:	Moral reasoning orientation polar scores for the 42 vision statements	53
Figure 7:	Moral reasoning orientation ratings of 42 vision statements on 4-point scale	55
Figure 8:	Moral reasoning orientation scores (polar) for vision statements aggregated by period: pre- or post-catastrophe	56
Figure 9:	Moral reasoning orientation ratings on 4-point scale for vision statements aggregated by period	57
Figure 10:	Explanatory style internality versus externality component for 42 vision statements	60
Figure 11:	Explanatory style stability versus instability component for 42 vision statements	61

Figure 12:	Explanatory style globality versus specificity component for 42 vision statements.....	61
Figure 13:	Explanatory style combined scores for 42 vision statements.....	62
Figure 14:	Explanatory style internality versus externality component for 42 vision statements aggregated by period: pre- and post-catastrophe.....	64
Figure 15:	Explanatory style stability versus instability component for 42 vision statements aggregated by period: pre- and post-catastrophe.....	65
Figure 16:	Explanatory style globality versus specificity component for 42 vision statements aggregated by period: pre- and post-catastrophe.....	65
Figure 17:	Explanatory style scores consolidating three components and preserving distinction between good and bad impact	66
Figure 18:	Frequency of ratings for internality versus externality component of explanatory style	67
Figure 19:	Percent rumination by vision: 1-3 before catastrophe and 4-6 afterward	69
Figure 20:	Percent rumination for vision statements aggregated by period: pre- or post-catastrophe.....	72
Figure 21:	Scatterplots of Moral Reasoning Stage (MS), Moral Reasoning Orientation – Caring (MOC), Moral Reasoning Orientation – Justice (MOJ), Explanatory Style – Good (ESG), Explanatory Style – Bad (ESB), and Rumination (Rum)	75
Figure 22:	Scatterplots of Rumination (Rum) versus Explanatory Style – Good (ESG), Explanatory Style – Bad (ESB), and Explanatory Style components (Good and Bad) Internality versus Externality (I/EG and I/EB), Stability versus Instability (S/IG and S/IB), and Globality versus Specificity (G/SG and G/SB)	77
Figure 23:	Moral reasoning orientation plotted on two axes (illustrative)	91

List of Appendixes

Appendix A:	Scoring Aid for Moral Reasoning Stage	108
Appendix B:	Scoring Aid for Moral Reasoning Orientation.....	113
Appendix C:	Scoring Aid for Explanatory Style.....	114
Appendix D:	Scoring Aid for Rumination.....	116
Appendix E:	Stages of Moral Reasoning Displayed Numerically	117
Appendix F:	Unweighted Moral Reasoning Orientation Ratings Displayed Numerically	120
Appendix G:	Explanatory Style Ratings Displayed Numerically.....	122
Appendix H:	Rumination Ratings Displayed Numerically	125
Appendix I:	Table of 42 Selected Vision Statements.....	127

CHAPTER ONE

Introduction

I could heartily wish this had not befall'n; but since it is as it is, mend it for your own good (Shakespeare, 1604/1997, p. 1266).

This study is about misfortune befall'n. It examines vision statements made before and after a catastrophe for changes beyond manifest surface meanings. The data for this study were the vision statements articulated by 7 pivotal national leaders before and after a human-caused catastrophe. Content analysis was used to search for latent changes in moral reasoning stage, moral reasoning orientation, explanatory style, and rumination. The outcomes were inferences about the effects of real-world catastrophes on pivotal leaders' vision statements. Table 1 shows the 7 leaders and the associated human-caused catastrophes selected for study.

"Crisis" describes a situation with such serious impact that a leader perceives the organization, or a significant portion of it, may not achieve fundamental goals or even survive (Nystrom & Starbuck, 1984; Weick, 2001). I use "catastrophe" to describe a crisis that is sudden. Nature can cause catastrophe, but this study focuses on catastrophes caused by people. A qualitative difference is that nature is not held culpable of wrongdoing in the way people who cause a catastrophe are held culpable. Human-caused catastrophes evoke qualitatively different feelings than natural catastrophes, perhaps due to people's unconscious or partially unconscious beliefs, perceptions, and thoughts (Schein, 1992).

Table 1

Pivotal Leaders and Catastrophes Selected for Study

Leader	Catastrophe	Catastrophe Date
George W. Bush	Terrorists attacked the United States	September 11, 2001
Margaret Thatcher	Argentina attacked the Falkland Islands	April 2, 1982
Golda Meir	Egypt and Syria attacked Israel	October 6, 1973
John F. Kennedy	Soviet Union placed offensive missile sites in Cuba	October 16, 1962
Franklin Roosevelt	Japan attacked Pearl Harbor	December 7, 1941
Woodrow Wilson	Germany waged unrestricted submarine warfare	April 12, 1917
Abraham Lincoln	Confederates captured Fort Sumter	April 14, 1861

While unconscious or partially unconscious beliefs, perceptions, and thoughts evoked by catastrophe affect both leaders and followers, this study focuses on the reactions of leaders through an examination of the vision statements leaders articulated proximately before and after a human-caused catastrophe. Psychological theory predicts there will be measurable differences in these texts (Winter & Stewart, 1977).

My search for usable vision statements identified a practical constraint: If leaders are categorized three ways—for profit, nonprofit, and public sector—sets of vision statements encompassing both sides of a catastrophe are visible only for public sector leaders. They may not exist for leaders in other categories for multiple reasons, for example the leader never voiced a series of visions, post-catastrophe statements were avoided for legal reasons, or vision statements were not documented. This study posits that how public sector leaders' vision statements change might clarify how the vision statements of any organization's leader do change or could change in response to a catastrophe.

Organizational development practitioners usually advise leaders to react to crisis by fostering adaptation, cooperation, trust and friendly relationships that steer followers away from defensive avoidance and toward confronting reality (Bass, 1998). Bernard Bass has written that any leadership theory of consequence implies that leaders use empathy to help followers adjust to environmental changes and that empathy is generally held to be "emotional, intuitive, and unconscious" (Bass, 1990, p. 123; Turner, Barling, & Zacharatos, 2002). However, measuring "empathy" has proved difficult, and Bass has called for more research that attempts to measure the construct. Empathy is just one of a number of emotions and motivations that are studied as prosocial behavior and a leader's vision statement is by definition prosocial, articulating an attractive improvement over peoples' perceived alternatives. So this study is responsive to Bass' call for new research designs to explore links between leadership and empathy and goes beyond that to explore links between leadership and the more inclusive concept of prosocial behavior.

Increasing Need for Catastrophe Leadership

Today's leaders must expect to handle an increasing number of human-caused catastrophes that touch large numbers of people. We live in an era characterized by strong tensions between global interdependence and expressions of diversity, by increased interaction across physical and political boundaries, and by closer technological and economic ties (Lipman-Blumen, 1996). These tensions are demonstrated by the emergence of global terrorism. Currently, we see extensive interactions across national boundaries that are focused on efforts to organize international collaboration to defend against global terrorism. Closer ties are demonstrated by the economic harm suffered by multiple nations due to a catastrophe like the terrorist attack on the United States on September 11, 2001, where the economic consequences had enormous impact abroad.

The need to handle more human-caused catastrophes also has fostered a new emphasis on catastrophe prevention, what leaders in earlier years referred to as

anticipating and avoiding problems. Some organizations have been characterized as having high risk of a catastrophe that would have exceptional consequences (e.g., nuclear aircraft carriers and nuclear power generating plants). Deliberate effort to manage such organizations as "high reliability organizations" is reported in the literature (Weick & Sutcliffe, 2001).

Vision Statement Criteria

Before 1985, scholars frequently used the term "purpose" to describe organizations' fundamental goals. Recently, the term "vision" has gained popularity because of its use by Bennis and Nanus. Vision statements are an observable artifact of leadership because leaders must repeat visions "time and time again" to reinforce them and incorporate them into the organization's culture (Bennis & Nanus, 1997, p. 101).

To be recognized as a vision, a leader's statement must meet these criteria: (a) accessible to substantially all organization members or their designated representatives (Fairhurst & Sarr, 1996; Kouzes & Posner, 1987), (b) an image of the future organization, and (c) an attractive improvement over perceived alternatives (Bennis & Nanus, 1997, p. 82; Burns, 1978; Deison, 1994; Kouzes & Posner, 1987; Nadler & Tushman, 1989). Although other criteria have been offered, they often prove unhelpful or sometimes wrong, for example "specific," "realistic," and "credible." The truth is, a useful vision statement may be quite the opposite: faith-based, values-based, belief-based, radical, bold and unconventional (Bass, 1998).

Problem Statement

My interest in this study arose at the time of the Los Angeles riots of 1992 when it became apparent that leaders need to learn how to deliver appropriate vision statements during times of catastrophe so everyone in the organization can do the right thing. The near-term practical solution in my organization was a Building Emergency Leader (BEL) System (Oliver, 1993) composed of specially selected and trained leaders linked by a disaster communication network.

This practical solution left a huge question unanswered however. What "taken-for-granted beliefs, perceptions, thoughts, and feelings" (Schein, 1992, p. 17) underlie vision statements that leaders use to guide their organization through a human-caused catastrophe? We need to look beyond the manifest content, the surface meanings, to latent content, deeper meaning embedded in the text (Holsti, 1969). This study posits that moral reasoning stage, moral reasoning orientation, explanatory style, and rumination are four measurable aspects of vision speeches that will reflect changes in latent content. This study undertakes to examine comparable sets of pivotal leaders' vision statements before and after a catastrophe to show correspondences or discrepancies in those four measures.

Key Definitions

Central to this study are the definitions for pivotal leaders, human-caused catastrophe, and vision statement. Other terms integral to the research relate to the measures used in the study and to supporting theories. The three central definitions are set forth below, followed by definitions of the other terms.

Pivotal Leaders

Formal leaders whose choices appear to set direction for their entire organization are pivotal leaders. A central assumption of this study is that what formal leaders say makes a difference for the organization. It is relatively easy to measure surface meanings. Various analysts and commentators do it every day. It is important also to measure deeper meanings latent in the leaders' statements that may provide insight into the beliefs, perceptions, thoughts and feelings that underlie those statements.

Vision Statement

A leader's prepared text is a vision statement if it meets these criteria: (a) accessible to substantially all organization members or their designated representatives, (b) an image of the future organization, (c) an attractive improvement over perceived alternatives. Vision statements are, or at least appear to be, key tools that leaders use to get everyone in their organization to advance in the "right" direction. It is important to measure the vision statements of pivotal leaders in the past to understand what has been effective through the years and distill principles that might be useful to leaders crafting new vision statements today and tomorrow.

Human-Caused Catastrophe

From the perspective of the leader, human-caused catastrophe is a sudden crisis with such serious impact that the organization, or a significant portion of it, may not achieve fundamental goals or even survive and that appears to have been caused by people. Much of the divisiveness and turmoil that has plagued the world through history is attributable to people who caused a catastrophe for other people, and when the catastrophe places survival of an organization or a nation at risk, leaders are expected to rise to the occasion and provide statements of vision and direction sufficient for the organization to overcome the danger. It is important to examine how leaders of the past have crafted vision statements to achieve such success.

Explanatory Style

Habitual patterns people use to explain uncontrollable good or bad events are their explanatory style (Peterson, 1992) and important because they may illuminate causal relationships: Why people behave as they do. Explanatory style is important to measure because leaders who understand what causes their own and others' behavior may be better able to guide that behavior to achieve the organization's goals.

Measures of Explanatory Style

Content Analysis of Verbatim Explanations (CAVE) is a method to assess peoples' own explanations of the causes of uncontrollable good or bad events. It is applied here to assess events that affect the pivotal leader or the pivotal leader's organization (Peterson, Buchanan, & Seligman, 1995; Zullo, 1988; Zullo & Seligman, 1990).

Moral Reasoning Stage

Moral reasoning stages are unique ways of thinking that evolve sequentially as a person ages and that govern the person's behavior consistently (Kohlberg & Armon, 1984). This is important as another approach to illuminating causal relationships: Why people behave as they do. It is important to measure because leaders who understand what causes their own and others' behavior may be better able to guide that behavior to achieve the organization's goals.

Measures of Moral Reasoning Stage

For this study, moral reasoning stage is measured by applying the Hierarchical Complexity Scoring System (HCSS) (Commons, Danaher, Miller, Goodheart, & Dawson, 2001; Commons & Richards, 2003). This measure has been further developed recently by construction of computer assessment software, the Lexical Abstraction Assessment System (LAAS) (Dawson & Wilson, in press).

Moral Reasoning Orientation

An individual's world view that structures thinking about the nature of moral conflicts, and which factors deserve priority when resolving them, is called moral reasoning orientation. It is important because two orientations that have been identified, justice or caring, may lead to very different resolutions of conflict. It is important to measure moral reasoning orientation because leaders who understand how people resolve moral conflicts may be better able to adapt their own behavior and influence others' behaviors to achieve the organization's goals.

Measures of Moral Reasoning Orientation

For this study, moral reasoning orientation is measured by applying Lyons' (1982, 1988) justice and caring scoring model.

Rumination

Rumination is thinking about bad events that have happened or are happening, the who, what, where, or how of a bad event, or explanation why a bad event occurred, or words expressing negative emotional state or display. Rumination may have important effects on a person's worldview and relationships. Rumination is independent of pessimism; a person can be a pessimistic ruminator or an optimistic ruminator. Rumination can be contagious and spread to others, or it can discourage support from

others (Zullo, 1995). Rumination is important to measure because leaders who understand how others in their organization perceive events may be better able to adjust their own and others' behavior to achieve the organization's goals.

Measures of Rumination

For this study, rumination is measured by applying Zullo's (1988) method of content analysis to assess percentage of rumination in a vision text.

Organizational Culture

The pattern of basic assumptions that organization members assume are truths based on unconscious beliefs, thoughts, and feelings constitute organizational culture (Schein, 1985, 1992). They cause organization members' behavior, and are important to measure because a leader who understands the culture may be able to control elements of it to better direct members' behavior toward an envisioned goal.

Prosocial Behavior

Actions intended to benefit people other than, or in addition to, one's self comprise prosocial behavior (Batson, 1998; Batson, Ahmad, Lishner, & Tsang, 2002). It is important because a leader who understands prosocial behavior may be better able to direct members' behavior toward an envisioned goal even if the members' personal goals are different.

Assumptions

This study assumes that moral reasoning stage, moral reasoning orientation, explanatory style, and rumination are variables and that it is important to study changes in these before and after a catastrophe occurred (Colby, Kohlberg, & Kauffman, 1987). A second assumption is that available texts of the leaders' vision statements are substantially accurate. To ensure high quality, texts will be obtained from official or authoritative sources of good reputation. While reality is that published texts do not always exactly match as-delivered texts, published texts are an approved version identified with the leader.

Research Question

In summary, the research question for this study is: What is the impact of human-caused catastrophe on pivotal national leaders' vision statements, as measured by correspondences and discrepancies in moral reasoning stage, moral reasoning orientation, explanatory style, and rumination before and after the catastrophe?

CHAPTER TWO

Review of the Literature

Introduction

In reviewing the literature, I undertook to answer five questions: (a) Does it matter that these national leaders' vision statements are influenced by advisors and consultants? (b) Does it matter what leadership theory lens is used for the research? (c) Do social psychology theories provide structural elements useful for this research? (d) Do developmental psychology theories provide structural elements useful for this research? (e) Are unobtrusive measures useful for this research?

I organized this review topically to focus on four literatures that I felt underlie this study. Broadly, they are leadership research, social psychology, developmental psychology, and non-reactive or unobtrusive measures.

With respect to leadership, I observed that most previous analysis was at the individual level but some has been at the organization level recently and the organizational approach might be particularly useful for this study. I also see value in studying real leaders in real-world settings to supplement experiments conducted in laboratories and other environments that are designed to limit and control variables. Content analysis techniques like those I use here have previously been applied to similar unobtrusive, at-a-distance assessment of real leaders in the real-world.

With respect to social psychology, I observed that streams of research into organization culture, prosocial behavior, and explanatory style might bear on this study. Edgar Schein's (1985, 1992) organizational culture research indicates how leaders, visions and organization members interact. Research into prosocial behavior supports examination of leaders' vision statements as artifacts with measurable prosocial elements. Explanatory style research originally focused on assessing individuals' optimism, pessimism, and depression but has evolved into a method for analyzing the tendency of leaders and their organizations to express views positively or negatively.

With respect to developmental psychology, previous research indicated that events and crises result in changes in the complexity of leaders' speeches. Moral reasoning stage and moral reasoning orientation are closely related concepts for which unobtrusive measures now exist to assess changes.

With respect to non-reactive or unobtrusive measures, because this study relies on them I briefly explore the literature on their strengths and weaknesses.

The Organizational Frame

At the national level, the vision statements of pivotal leaders are organizational culture artifacts created by multiple people (Frum, 2003; Gelderman, 1997; Ritter & Medhurst, 2003). Winter and Stewart (1977) underscore this reality:

[T]he measures used here reflect not only the motives of the president, but also those of his inner circle of advisors—that is, the "atmosphere" of the administration—this should not rule out their predictive value, since many presidential decisions and behaviors are also likely to be the products of this same circle. (p. 50)

A review of the literature suggests the level of analysis researchers use makes a substantial difference. If analysis is at the individual level, seeking to predict an individual's behavior or seeking to interpret an individual's personality, the influence of advisors and consultants is a confounding factor that degrades interpretations and inferences. If, however, analysis is done at the organization level and assumes the organization's leadership is composed of a formal leader working together with advisors and consultants, then the multiple influences on vision statements are not confounding factors but are part of the integrated phenomenon under study (Dille, 2000; Walker & Schafer, 2000). Schafer (2000) addressed this issue specifically:

Rarely does a state leader make a decision in isolation without at least some input from others in the process. In fact, it may be the case that our models need to be concerned with some form of aggregation of psychological characteristics to best understand the policies of the state. (p. 513)

Similarly, Axelrod (1976) cautioned against seeing analysis as revealing truth about the speaker's inner self rather than image presented by the speaker. So preliminary analysis suggests that level of analysis is important: It should be at the organizational, not individual, level.

Leadership Research

Origin and Trend

The history of organized leadership research in the United States often is analyzed as a time sequence of four approaches dating from the early 1900s. Until well into the 1940s, research emphasized traits that might mark individuals as leaders, especially traits that might distinguish leaders from followers. From the 1940s to the 1960s, the literature focused on behaviors marking individuals as leaders. From the 1960s to the 1980s, the research efforts highlighted how leaders and situations mediate each other; and since the 1980s, researchers have grappled with how individuals as leaders manage meaning. A fifth approach has emerged recently: the study of leadership effected by multiple people, dispersed or shared leadership (Bryman, 1996; Pearce & Conger, 2003a).

Together, these four or five lenses are perceived as complementary (Bass, 1990; Bryman, 1996; Pearce & Conger, 2003a; Stogdill, 1974). The first four lenses require analysis at the level of the individual as a leader. For the fifth approach, the level of

analysis must be larger: partnerships by "ones and twos" at the top of corporations (Bennis, 1999; Heenan & Bennis, 1999), team members (Neubert, 1999), work group, department, or organization (George & Bettenhausen, 1990).

Salient Research

Consistent with these categories, but with a selective eye, four paradigms of leadership research might be called analysis of great leaders, experimental observation of leader behaviors, at-a-distance assessment of leaders, and analysis of shared leadership.

Analysis of great leaders. A strength of these early leadership studies was their focus on leaders operating in the real-world. The writings of Plato, Machiavelli, and Confucius, for example, distill the principles of leadership as seen from their respective cultures in the context of their perceptions of history, human nature, and trends in world events. This approach continues to be popular and generates biographies and studies such as James Bruce's (1986) study of 13 chief executive officers.

Criticism leveled at these studies has suggested various "vagaries and imprecision" of method (Chemers, 1997, p. 18). Leaders are often selected more subjectively than objectively. Data are collected by methods of convenience, and analysis is based on emergence of themes recognized intuitively. On the other hand, these studies describe what real-world leaders and their commentators perceive is responsible for their successes. They are valuable as case studies: leaders studied for what they personally have done and not as samples representing some population called leaders (Robson, 1993).

Great leader studies demonstrate real-world research (Robson, 1993), where the investigator has little or no access to the leader and no control over variables (Stogdill, 1974). Real-world research is an earnest effort "to say something sensible about a complex, relatively poorly controlled and generally 'messy' situation" (Robson, 1993, p. 3). The critics' complaints encourage investigators undertaking real-world research to maximize rigor in their methods so research outcomes are credible, independently reproducible, and offer comparison of the same leader at multiple times or comparison of multiple leaders (Robson, 1993; Winter & Stewart, 1977).

Experimental observation of leader behaviors. To obtain the advantages of greater rigor, many researchers in the 20th century carried out experiments in laboratories or other environments designed to limit and control variables. When preparing the first edition of the *Handbook of Leadership*, Ralph Stogdill deliberately biased the work toward experimental studies he found "competently executed" (Stogdill, 1974, p. viii). Two subsequent editions of the handbook, authored by Bernard Bass, continued the bias toward experimental studies but deliberately accepted studies with less rigor on grounds that "much of the understanding about leadership occurs from testable ideas that have not been, as yet, fully tested" (Bass, 1990, p. xv).

The advantage of experimental method is that significant control over variables permits careful design and execution that improve independent reproducibility and

credible comparisons of the same leader at multiple times or comparisons of multiple leaders. Experiments may even yield generalizable truths. On the other hand, control over variables often makes the research environment so unrealistic that outcomes are not meaningful in the real world. When preparing the third edition of the *Handbook of Leadership*, Bass (1990) observed that experiments involving college-age participants in laboratory settings had decreased while studies of working leaders in their field environments had increased.

At-a-distance assessment of leaders. Another stream of research began in earnest during World War II and has largely been sponsored by the U.S. government through the Central Intelligence Agency. A few pre-World War II studies are cited as even earlier examples, including a 1930s¹ study of Woodrow Wilson by S. Freud and W. C. Bullitt (1999) and John McDiarmid's 1937 study of American presidents' inaugural speeches. The watershed studies were two at-a-distance assessments of Adolf Hitler, the first a descriptive study completed in December 1942, and the second a more analytical study completed in December 1943 (Langer, 1972; Post, 2003a). Beginning in the 1960s, the work evolved from a sporadic effort into a U.S. government sponsored program that systematically assessed high interest leaders and offered a standardized conceptual framework for an integrated personality profile (Post, 2003a, 2003b). Finished reports read much like a clinical psychologist's description and diagnosis of a patient (e.g., William Clinton and Saddam Hussein in Post, 2003b).

Of relevance to this study, at-a-distance leader assessments have used content analysis of leaders' texts to evaluate motives, traits, and cognitions. These analyses were deliberately rigorous. "Typically, these measures are carefully designed, with examples and training procedures, to enable previously inexperienced scorers to apply them with high reliability" (Winter, 2003a, p. 22). The measures used have included motive imagery (Winter, 1991), operational codes (Schafer, 2000), integrative complexity (Suedfeld & Tetlock, 1977; Tetlock, 1979), cognitive mapping (Axelrod, 1976; Hart, 1977), and explanatory style (Satterfield, 1998; Satterfield & Seligman, 1994; Zullow, Oettingen, Peterson, & Seligman, 1988).

The strengths of this approach include the study of real leaders working in real-world situations, both living leaders and leaders of yesteryear. This includes leaders working under conditions so urgent that realistically they are not available for laboratory testing or, if they were, ethical constraints would preclude public disclosure of results (Winter, 2003a).

A principal criticism of at-a-distance research has been the issue of who is being analyzed, the leader or speechwriters and advisors. For some cultures, it appears well established that a prominent leader "merely delivers utterances written by functionaries (e.g., the British and Canadian Speeches from the Throne and the speeches of Japanese prime ministers)" (Suedfeld, Guttieri, & Tetlock, 2003, p. 266). One at-a-distance

¹ Bullitt reported the research was conducted over 10 years, signed in 1932, changed in 1938, and first published in 1966.

researcher responds to this criticism by arguing that scoring of personal letters and public statements of the same leader uncover no significant differences on measures of integrative complexity (Suedfeld et al., 2003). Leaders choose assistants able to create texts that suit the leader's purpose, and leaders review and rework the text until it is what the leader wants (Winter, 2002, 2003b). The criticism is met by changing the level of analysis from individual to organization, from the individual-as-a-leader to the leader-as-a-construct. Multiple influences are no longer confounding but instead become part of the integrated phenomenon under study.

Analysis of shared leadership. The fourth relevant stream of literature, analysis of shared leadership, found organized expression beginning in the 1990s, although elements existed earlier in the literature (Bryman, 1996; Pearce & Conger, 2003a). Shared leadership is defined as "a dynamic, interactive influence process among individuals in groups for which the objective is to lead one another to the achievement of group or organizational goals or both" (Pearce & Conger, 2003b, p. 1). For example, Leonard Berkowitz (1953) explored participants' reactions when conference leaders allowed others to share control over meeting process. Richard Hodgson, Daniel Levinson, and Abraham Zaleznik (1965) studied how three psychiatrist-executives cooperatively supervised a mental hospital. David Bowers and Stanley Seashore (1966) analyzed agents' reports of leadership acts by both supervisors and peers in the insurance industry. David Heenan and Warren Bennis (1999) studied cooperation between Chief Executive Officers (CEOs) and their deputies at corporate headquarters (Bennis, 1999; O'Toole, Galbraith, & Lawler, 2003).

A strength of shared leadership research is its recognition that even the most pivotal leader in an organization—perhaps especially the most pivotal leader—is unlikely to act alone. O'Toole et al. (2003) argued that "even the most fabled 'solitary' leaders relied on the support of a team of other effective leaders" (p. 251). They listed, as examples, celebrity leaders like George Washington, Franklin Roosevelt, and Winston Churchill.

A criticism of shared leadership research focuses on practicalities in the locus of power. One practical consideration is cultural: Western culture holds a single person accountable for the actions of an organization (O'Toole et al., 2003). Another practical consideration is pecking order: Edwin Locke (2003) opined that shared leadership would be very fragile if it exists, "It is almost inevitable that the people at the top will disagree on some, if not many, issues, and it would paralyze decision making if neither had the final say" (p. 276).

As early as the first edition of the *Handbook of Leadership*, Stogdill (1974) addressed the distribution and sharing of leadership power, but criticism focused on power issues may miss the point. The definition of shared leadership posits "dynamic, interactive influence process" (Pearce & Conger, 2003b, p. 1) and is silent on the issue of power. All of the shared leadership research examples cited above focus on the dynamic and mutual interactive influence of two or more people working together. Equality of power never is asserted. What remains unclear is how to objectively

distinguish between leader-follower interaction and shared leadership. Seibert, Sparrowe, and Liden (2003) used leader-member exchange (LMX) theory to study this issue but the results were inconclusive.

Application and Measure

While other leadership research approaches generally assume that a leader is an individual and focus analysis at the individual level, the shared leadership approach generally assumes that multiple people in the group perform leadership tasks and focuses analysis at the group level.

With respect to vision statements by pivotal leaders, evidence shows that people other than the leader do influence vision statements (Gelderman, 1997; Ritter & Medhurst, 2003; Suedfeld et al., 2003; Winter, 2003b), but it is difficult to discern precisely whether that influence is large (Kennan & Hadley, 1986) or small (National Archives and Records Administration, 2001). At the national level in the United States and elsewhere, pivotal leaders do seek assistance in crafting their vision statement speeches. Medhurst (2003) lists some of the many advisors who helped U.S. presidents: Alexander Hamilton, James Madison, and John Jay assisted George Washington; William Seward influenced Abraham Lincoln; Raymond Moley, Rexford Tugwell, Thomas Corcoran, Benjamin Cohen, Samuel Rosenman, Robert Sherwood, and Archibald McLeish were aides to Franklin Delano Roosevelt; Ted Sorensen helped John Kennedy, and so forth to current times. The help involved articulation and writing of ideas and sometimes creation of ideas and policy but the mutual interaction theme emerges:

[T]here is simply no truth to the charge that words are being put into the president's mouth. The presidents are too involved and the staffing process is too rigorous for anything like that to happen. The typical presidential speech . . . is vetted by anywhere from 5 to 20 people before it is finalized (p. 9).

Medhurst (2003) posited that extensive group input means speeches do not voice just personal ideas of the president but instead represent the nation and its people. Eight other scholars provide evidence supporting Medhurst's view in the same volume (Ritter & Medhurst, 2003). In describing personal experience as one of George W. Bush's speechwriters, David Frum (2003) reported the speechwriting process included recognition of what the leader wants to say (requirements), what the leader has said in the past (policy and precedent), and editing (tailoring). Many people contribute ideas and words to the vision voiced by the "formally acknowledged leader" (Avolio, Jung, Murry, & Sivasubramaniam, 1996).

While the influence of any one person, formal leader or advisor, is difficult to measure and may vary from one vision statement to another, available descriptions suggest the process of crafting visions has changed from creation by the leader, or by the leader and an aide, to a synthesis of dynamic inputs from an interactive team, at least for

leaders operating in a democratic, elective culture. Study of the presidential speechwriting process has spawned a number of books. Here are selected summary descriptions of the process relevant to this research:

Lincoln. William H. Seward, Secretary of State, is known to have written a rough draft of the last paragraph of Lincoln's first inaugural address (Podell & Anzovin, 2001) and to have contributed suggestions and changes to Lincoln's message to Congress delivered on July 4, 1861 (Basler, 1953).

Wilson. Wilson, seemingly alone as was his custom, wrote his war message to Congress during the early hours of April 1, 1917 (Freud & Bullitt, 1999). Wilson was the first president since John Adams to personally speak to Congress, a technique he used to lead public opinion and continue a movement begun by Theodore Roosevelt intended to change the president from an administrator to an influential national voice—the only voice elected by all voters in the nation (Gelderman, 1997).

Franklin Roosevelt. To prepare speeches, Roosevelt routinely met with several aides after dinner and together they dictated a draft and revisions. Two or three aides, usually Samuel Rosenman, Robert Sherwood, and Harry Hopkins, continued working through the night to produce another draft that was placed on Roosevelt's breakfast tray. This process continued for several days, through a dozen or more drafts. By the time a final reading copy was produced, Roosevelt knew it almost by heart because he had studied it, edited it, added to and deleted from it, and critiqued it (Gelderman, 1997).

Kennedy. Kennedy relied on Theodore Sorenson to draft speeches. Kennedy's process included convening a group of advisors and listening to their diverse arguments to gather information useful for writing the speech (Windt, 2003).

Thatcher. Thatcher worked extensively with speechwriters, arguing, criticizing, and rejecting draft after draft. She insisted that every word go through "the mincing machine of my criticism" to ensure she knew the facts and understood the arguments (Campbell, 2003, p. 32).

Bush. Speechwriters for Bush begin with instructions received directly from the president. Bush edits drafts severely, insists on "strict linear logic," and requires statistics although he rarely speaks them. Many writers contribute suggestions, then a lead writer integrates them into a coherent whole (Frum, 2003, pp. 48, 141, 145, 147). At a press conference, Bush acknowledged that others provide information for his speeches and said, "I take personal responsibility for everything I say . . . Absolutely" (Bush, 2003).

The vision statements of pivotal national leaders build on ideas intended to represent organization members' interests and are often vetted by 5 to 20 people working together to help the organization's formal leader achieve the organization's goals. They do not represent the leader-as-an-individual but do represent shared leadership.

This study views pivotal national leadership as shared leadership composed of a formal leader working together with advisors and consultants. National vision statements are created and vetted by multiple people. There is tension in this sharing. In our culture, formal leaders seek to benefit from the collective wisdom of selected and trusted advisors but recognize that they will be held personally accountable for the vision they

speak. Of course, vision statements are presented as the formal leader's words. They have effects as the formal leader's words (Winter, 1992, 2003b), and the formal leader's reputation stands or falls on those words (Nanus, 1992). President John F. Kennedy made the point: "the President bears the burden of responsibility . . . the advisors may move on to new advice" (Bruce, 1986, p. 21).

So the answer to the first of my questions is that the level of analysis is important. It needs to be at the organizational level. It is logical for the formal leader of an organization to work with advisors and consultants and their influence needs to be an integral part of what is being studied. The answer to the second of my questions is that the shared leadership approach fits organization level analysis and examines the effects of multiple people working leadership together.

Social Psychology

Question three asked if social psychology provides structural elements useful for this research. From social psychology, three relevant streams of research include organizational culture, prosocial behavior, and explanatory style.

Origin and Trend: Organizational Culture

Edgar Schein proposed that organizational culture exists at three levels. At the surface are artifacts: objects, processes, and structures an observer can readily see. Those artifacts can be traced to values espoused by the organization and its members: their philosophies, their goals, and their strategies. In turn, those values arise from basic underlying assumptions that are hard to recognize: unconscious beliefs, thoughts, and feelings that organization members assume are truths. Schein offered an analysis of artifacts, values, and basic assumptions in his research at the Action Company and the Multi Company (Schein, 1992).

Salient Research: Organizational Culture

Schein's lens is only one of many used to study organizational culture and social psychologists have not reached agreement on what approach should be used. Martin and Frost (1996) chose a war metaphor to characterize the diversity of organizational culture approaches: lenses contending for acceptance include rational theory, function theory, symbolics, integration, differentiation, fragmentation, clinical theory, critical theory, stage change theory, mosaic, nexus, and meta-theory. In organizational culture research, said Martin and Frost, "there is chaos rather than order" (p. 599); however, Schein's lens has face validity for this study.

Application and Measure: Organizational Culture

Schein (1992) advocated encouraging input from multiple members of the organization. He believed that organizations can self-organize, diagnose needs, and implement culture changes, and his method involves organization members in a joint search for data and solutions.

With respect to vision statements, Schein (1992) has recognized two forms: strong or fuzzy. Where the future is reasonably predictable, a vision can be strong: There is a clear goal and a path to get there. Where the future is uncertain, a vision at start can be fuzzy: Change is set as a goal, but other organization members are expected to provide expert input. "In our obsession with leadership vision, we may have made it possible for learning leaders to admit that their vision is not clear and that the whole organization will have to learn together" (p. 383).

Schein (1992) also offered the insight that the effectiveness of organizational vision is mediated by situation. Followers who do not see the need to change are unlikely to pay much attention to a leader's vision. "[N]ew visions are most important when people are ready to pay attention, and they are only ready to pay attention when they are consciously or unconsciously hurting because of an accumulation of disconfirming information" (p. 301). Schein's statement illuminates why rumination is an appropriate measure for this study. By definition, rumination is the internal articulation of disconfirming information—thinking about bad events that have happened or are happening, the who, what, where, or how of a bad event, or explanation why a bad event occurred, or words expressing negative emotional state or display.

Satterfield (1998) explored the relationship between rumination and action among political leaders with respect to significant military and political events. Pessimistic rumination scores did not predict either aggression or risk taking. He opined that for his purpose scoring any negative emotion, event, or explanation was too broad; he needed a measure that would focus on doubts and problems that might cause the leader to feel less confident and perhaps disengage from further action. Similarly, in this study rumination scores are used to gauge, at the organization level, the leader's articulated reactions to crisis. In Schein's (1992) terms, a leader's ruminative reactions may influence followers to pay attention to the leader's vision statement.

Origin and Trend: Prosocial Behavior

Prosocial behavior, the opposite of antisocial, is action intended to benefit people other than, or in addition to, one's self. Altruism, collectivism, and principlism are thought to motivate prosocial behavior. Altruism is a desire to benefit another person, collectivism is a desire to benefit a particular group of people, and principlism is the desire to uphold a universal moral principle, for example, justice (Batson, 1998; Batson et al., 2002). This illuminates why moral reasoning orientation is an appropriate measure for this study: Justice orientation upholds universal moral principles (principlism) and caring orientation seeks benefit to another person or group (altruism or collectivism).

One criterion for vision statements is that they offer organization members an attractive improvement over perceived alternatives. A pivotal leader's vision statement does envision benefit to another person or a group or the upholding of a moral principle, so whether motivated by altruism, collectivism, or principlism, a vision statement is prosocial.

Salient Research: Prosocial Behavior

Substantial research into prosocial behavior followed a New York City incident in 1964: the murder of Kitty Genovese. A *New York Times* article reported Kitty was stabbed on a city street while 38 witnesses heard her screams and pleas for help but did not intervene or call police. Some question the accuracy of the *New York Times* article; nonetheless it launched numerous studies of prosocial behavior by witnesses to an emergency.

Early investigators included Darley and Latané (1968), who studied how individuals behaved when they heard a person nearby suffer an epileptic seizure. Latané and Darley (1968) looked at what happened when people found smoke filling a room. Amato (1983) studied the reaction of individuals who encountered one of five situations: a student asking their favorite color, a pedestrian falling with an injured leg, a pedestrian dropping 20 envelopes, a sidewalk salesperson, and one stranger overheard giving incorrect directions to another stranger. In another approach, investigators tested whether individuals who found apparently lost letters would mail them when they knew the letters contained information favoring or disfavoring controversial topics, such as "friends of the Nazi party," pro- or anti-civil rights groups, election candidates, and Chinese and Vietnamese attitudes (Milgram, 1969).

In the stream of research into prosocial behavior, the primary focus has been individuals' responses to dramatic stimuli and analysis has aggregated individuals' responses. This study differs in making group-level behavior the primary object of study since we are considering leaders as constructs. This study posits that in vision statements moral reasoning stage, moral reasoning orientation, explanatory style, and rumination are latent content that measure elements of prosocial behavior. Some of those four elements correspond to characteristics other researchers associated with prosocial behavior. Principled moral reasoning, measured in this study by moral reasoning stage and moral reasoning orientation, has been associated with altruism (Batson, 1998; Batson et al., 2002; Dozier & Miceli, 1985). Outrage, guilt, and shame are expected to emerge as rumination in this study, and all three have been associated with principledness (Batson, 1998).

Origin and Trend: Explanatory Style

Explanatory style research usually is traced to animal learning experiments that exposed dogs to electric shocks they could neither avoid nor stop. When the dogs later were exposed to electric shocks they could easily learn to stop, they did not even try but instead passively endured the shocks: they had learned to be helpless in this situation. This model was perceived as too simplistic to explain human behavior and the difference was cognitive: When people encounter an uncontrollable event, individual differences influence their explanation of why it happened and their perception of their personal control over the situation. These differences shaped the attributional reformulation of how learned helplessness works (Peterson et al., 1995).

The evolution of explanatory style research proceeded down a clinical path that explained behavior in degrees of optimism or pessimism and depression. The initial measure for explanatory style was the Attributional Style Questionnaire (ASQ). The ASQ assumed that people who habitually attributed uncontrollable bad events to stable (always present) causes internal to themselves were more likely to present symptoms of depression than people who oppositely attributed such events to unstable (one-time) causes external to themselves. The ASQ also measured how global such attributions are: Global attributions affect more of a person's life than specific ones do (Peterson et al., 1995). These three dimensions became the explanatory style standard: internal versus external, stable versus unstable, and global versus specific.

Development of a non-reactive, two-stage content analysis method called Content Analysis of Verbatim Explanations (CAVE) provided a reliable, consistent and valid measure of explanatory style and resulted in two changes to the research. First, CAVE created opportunity to measure explanatory style at a distance—over time, geography and the like (Peterson et al., 1995). Second, CAVE created opportunity to change the level of analysis from individuals to groups. Where research once studied individuals' binge eating or compulsive gambling (Peterson et al., 1995), new studies examined group-level phenomena: basketball and baseball teams' performance (Rettew & Reivich, 1995), late 19th century Russian Jewish and Orthodox Christian cultures, and cultural differences between East and West Berlin (Oettingen, 1995).

The extension of explanatory style to group-level analysis has included studies of leaders' speeches (Zullow, 1995). Zullow and others have used explanatory style to describe candidates for election to the U.S. Presidency or the U.S. Senate as well as President Lyndon B. Johnson's statements at press conferences—all examples of measuring leadership as a construct. Indeed, staff members for one presidential candidate, Michael Dukakis, reportedly used explanatory style to shape his nomination acceptance speech (Seligman, 1998; Zullow, 1995; Zullow et al., 1988).

Extending explanatory style from individual to group analysis has caused changes in operational definitions. Originally, text to be rated for explanatory style was extracted only if the event described affected the "speaker" personally (Peterson, Schulman, Castellon, & Seligman, 1992). Later the criterion was changed to allow extractions that describe events affecting one's self, one's in-group or political party, or one's nation (Zullow, 1988; Zullow & Seligman, 1990). A second change implied by Zullow and Seligman (1990) was redefinition of explanatory style as "the optimism or pessimism with which we explain the causes of bad events in our life" (p. 53). This focus on bad events exclusively appears to have been transitory; in a 1995 book, *Explanatory Style*, edited by Buchanan and Seligman, investigators were encouraged to expect a "dearth of positive events in verbatim materials" but not feel deterred in searching for them (Reivich, 1995, p. 32), and the 2002 *Handbook of Positive Psychology* devoted a chapter to "Optimistic Explanatory Style" (Snyder & Lopez, 2002).

Salient Research: Explanatory Style

Candidates who became President of the U.S. from 1948 to 1984 delivered nomination acceptance speeches with more positive (or less negative)² explanatory styles and with lower rumination scores than the opponent's 90% of the time. For 1900 to 1944, the same result occurred 75% of the time. The three exceptions were re-elections of Franklin Delano Roosevelt. The investigators opined that environmental influences may have affected those elections because Roosevelt's presidential era, from start to finish, faced severe crises including economic depression and global war. In Roosevelt's case, "ruminating about problems may have been necessary" (Zullow, 1995, p. 194).

Candidates elected also were more positive (or less negative) in nomination acceptance speeches and when seeking reelection than in their inauguration speeches (Zullow, 1995; Zullow & Seligman, 1990). Satterfield and Seligman (1994) assessed verbatim materials, primarily press conferences and interviews, created by Saddam Hussein and George H. W. Bush just before significant military or political events. They found optimistic explanatory style (positive scores) before aggressive and risky behavior and pessimistic explanatory style (negative scores) before cautious and passive behavior.

Application and Measures: Explanatory Style

The terms optimism, pessimism, and depression require further definition as reference points for explanatory style measurements. The terms originally were adopted in an effort to provide meaningful explanations of findings that the psychological community and the general public could both understand. "Explanatory style does not necessarily mean anything to most people, but an optimistic view of the causes of events certainly does" (Peterson et al., 1995, p. 10). It is important to note that most findings are not at the extremes, where individuals might be categorized as optimistic or pessimistic or depressed, and those labels do not apply well to organizational-level analysis. For this study I have chosen to substitute positive for optimistic and negative for pessimistic as words more accurate and meaningful for the organizational-level analysis being undertaken.

Explanatory style measures how people explain the cause of uncontrollable good or bad events that affect them. It has been measured on 7-point scales on three dimensions: the degree to which the explanation identifies the cause as (a) internal versus external ("I did it" vs. "They did it"); (b) stable versus unstable ("The cause will always exist" vs. "The cause was momentary"); and (c) global versus specific ("It affects everything" vs. "It only affects one thing"). Researchers usually have added the scores of these three dimensions into a single, composite score.

² The reference labels "pessimism" and "optimism" usually associated with explanatory style analysis at individual level are replaced for this study by the terms "negative" and "positive" that are more meaningful for analysis at the organizational level.

Buchanan and Seligman (1995) have called for more investigations that analyze the dimensions separately and compare them to the composite. Such studies are bringing change to explanatory style research. Recently, Peterson (2000) reported: "the internality dimension has become of less interest . . . [it] may well conflate self-blame and self-efficacy" and he suggested that a more valid measure may emphasize only stability and globality. Because of current interest in all of the scores, the three dimensions as well as the composite, this study will report and compare them all.

The measurement of explanatory style requires that good and bad events be considered separately; attributions for good and bad events frequently show independence rather than correlation (Peterson et al., 1995). This study adopts the modifications others have used to apply explanatory style to groups. Extractions may be selected when the event affects "my country," "my organization," or "myself." As noted above, one additional modification is to substitute positive for optimistic and negative for pessimistic as scalar directions.

So, the answer to my question three is that three streams of social psychology that provide useful structural elements for this research are organizational culture, prosocial behavior, and explanatory style. Organizational culture identifies vision statements as cultural artifacts perhaps important to followers who are ready to pay attention because they are ruminating about bad occurrences. Prosocial theory identifies vision statements as inherently prosocial because they intend to benefit people other than, or in addition to, the leader. Explanatory style presents a method of non-reactive analysis, CAVE, to assess leaders' statements about bad or good occurrences.

Developmental Psychology

Question four asked if developmental psychology provides structural elements useful for this research. Relevant streams of research from developmental psychology involve moral reasoning. The fundamental stream, moral reasoning stage, originated in the work of Piaget and was most influentially advanced by Kohlberg. One important offshoot is a focus on alternative modes of moral reasoning orientation posited by Gilligan.

Origin and Trend: Moral Reasoning Stage

Piaget's stage theory. Through observation and interview of children performing standardized tasks that he assigned, Piaget became convinced that intellectual development in children proceeds along a single, unvarying path that includes moral reasoning (Batson, 1998; Commons et al., 2001; Piaget, 1976). He described this path in terms of four stages, each composed of three parts: two sub-stages and one consolidated stage. Mathematically, Piaget's stages can define 12 levels of moral reasoning and Commons and colleagues argued that all 12 should be considered hard stages (Commons et al., 2001; Commons, Trudeau, Stein, Richards, & Krause, 1998; Dawson, Commons, & Wilson, 2001). In brief, a hard stage is identified by four criteria: a unique way of thinking, that evolved sequentially, governs the person's behavior consistently, and that

replaced prior stages by adding to the thought structure and transforming it to be greater than simply the sum of its parts (Kohlberg & Armon, 1984).

A stage, as envisioned by Piaget and later by Kohlberg, is a hierarchical structure that organizes and transforms the elements of the next lower stage to produce new, non-arbitrary, combined actions. Commons and associates described orders of hierarchical complexity, with a sequence of developmental stages that rank actions from least complex to most complex, the Model of Hierarchical Complexity (Commons et al., 1998). This was the foundation for the Hierarchical Complexity Scoring System (HCSS) structured around definitions of 15 hierarchical stages (Commons et al., 2001; Commons & Richards, 2003). Subsequent development by Dawson and Wilson (in press) redefined the same stages into 22 computer-identifiable stages.

Kohlberg's standardized dilemmas. Kohlberg undertook a dissertation extending Piaget's research to the moral judgment of adolescents age 10 to 16 but decided the data were better explained by a six-stage developmental typology. Kohlberg drew on the work of Baldwin and Kant as well to define six stages of progression in moral reasoning (Tappan, Kohlberg, Schrader, Higgins, Armon, and Lei, 1987). Kohlberg's Standard Issue Scoring System formally described each stage in terms of how people think about rights, justice, rules, laws, and punishment. Stage 1 actions avoid punishment (People caught stealing drugs go to jail). Stage 2 actions benefit the actor personally (It's OK for someone poor to steal drugs from the rich). Stage 3 actions conform to the mores of the group (A person who always helps everyone should not be punished for stealing drugs he needs). Stage 4 actions support a logical system or society (Anyone who steals drugs must expect courts to punish him to the degree of the just debt he owes society). Stage 5 actions benefit society by using rational principles (One who stole drugs should not be punished if police tricked him into confessing). Stage 6 actions act in accordance with universal principles of justice (It is always right to steal a drug if that saves a human life) (Colby, Kohlberg, & Kauffman, 1987).

Salient Research: Moral Reasoning Stage

Kohlberg's data showed positive correlation between age and moral reasoning stage "as one would expect of a developmental variable" (Colby, Kohlberg, Gibbs, & Lieberman, 1987, p. 100). The primary change over the years was a reduction in the number of participants at the lower stages (1, 2, and 3) and an increase in the number of participants at Stage 4. Numbers at Stage 5 increased slightly after age 20-22 but always remained small. Stage 6 never appeared and Kohlberg's scoring manual did not attempt to define it nor had its existence been validated. The data showed moderate positive correlation between education and moral reasoning stage and suggested that attending college facilitated achieving Stage 4 and completing college facilitated transition to Stage 5.

Kohlberg was criticized for changing his stage definitions and scoring methods. He defended every change as an application of lessons learned through research, a process he identified as "bootstrapping," and he asserted that the improvements reduced

subjective judgments raters were expected to make and increased the objective application of scoring exemplars (Colby, Kohlberg, & Kauffman, 1987).

A widely recognized weakness of Kohlberg's system is that it is domain specific. Kohlberg built the system around a one-on-one reactive interview, developed in 1958, that presents three standardized hypothetical moral dilemmas to each participant and elicits responses about paired moral issues: life versus law, conscience versus punishment, contract versus authority. Responses are scored by comparing them to specimen responses published in a scoring manual (Colby, Kohlberg, & Kauffman, 1987). The interview method, the prescribed content of the hypothetical dilemmas, and the match to specimen responses all constrain evaluation of moral reasoning stage to people accessible for interview and to the specific topics selected for dilemmas and specimen responses.

Less widely recognized, evidence suggests that Kohlberg's lower stages underestimate the ability of children under age 10 to see others' viewpoints. A possible explanation is that the sample population Kohlberg used to construct specimen responses contained no participants younger than age 10 (Dawson & Kay, 2001).

Kohlberg's system is grounded in a justice and rights moral reasoning perspective which seemed masculine to some women researchers. Carol Gilligan (1988a, 1988b, 1993; Gilligan & Attanucci, 1988) documented the existence of moral reasoning grounded in an alternative perspective, caring and response. While Gilligan's research population was women, she opined that men also could use a caring and response moral orientation.

Additional research did not support the idea that gender differences exist in moral orientation (Colby, Kohlberg, & Kauffman, 1987; Stevens-Long & Commons, 1992) and Johnston (1988) determined that adolescents routinely can access both justice and caring perspectives and can choose to use either. Kohlberg remained unpersuaded that caring constituted a separate moral domain inadequately addressed by his system. He asserted that even if the dilemmas posed conflicts of rights and justice, participants were free to respond from a caring and response perspective (Colby, Kohlberg, & Kauffman, 1987).

Application and Measures: Moral Reasoning Stage

In social psychology, the investigation of personality has generated theory about conceptual complexity as a personality trait. People rated low on conceptual complexity tend to perceive situations dichotomously or simplistically and people rated high tend to simultaneously recognize multiple perspectives that can be ambiguous or in conflict (Dille & Young, 2000). From conceptual complexity, research into decision-making generated theory about integrative complexity as a variable describing an individual's information processing capabilities. Integrative complexity has three components: (a) discrimination (ability to distinguish one thing from others), (b) differentiation (ability to simultaneously recognize multiple characteristics of a thing), (c) integration (ability to connect things together in many ways) (Suedfeld, Tetlock, & Ramirez, 1977).

In developmental psychology, complexity is envisioned as change over the life span in the direction of greater ability to integrate accumulated experiences of a lifetime (Stevens-Long & Commons, 1992). Piaget believed cognitive complexity, as reflected in moral reasoning, progressed in hierarchical stages.

A significant evolution in moral reasoning stage theory evolved about 1978 to 1980 when systematic and metasytematic stages were formulated by Commons and Richards (Richards & Commons, 1984; Commons & Richards, 2003) and on a different track when Fischer (1980) developed skill levels. Commons and Richards drew on Inhelder and Piaget (M. L. Commons, personal communication, December 6, 2003; Commons et al., 1998) while Fischer drew from the work of Pascual-Leone (e.g., Pascual-Leone, 1970). Fischer predicted stages higher than Piaget and Kohlberg had found, and Commons and Richards showed those stages in 1978 (M. L. Commons, personal communications, June 9 and December 6, 2003; Commons & Richards, 2003).

Compare Piaget's four basic stages to the 15 stages of Michael Commons' Model of Hierarchical Complexity (MHC). Piaget's sensorimotor stage was divided into four stages by MHC (calculatory, sensory and motor actions, circular sensory-motor actions, sensory-motor). With some overlap, Piaget's early pre-operational stage was divided into two stages (nominal and sentential). Piaget's pre-operational stage was retained (pre-operational), and Piaget's early concrete and concrete stages were retained (primary and concrete). Piaget's early formal and formal operational stages were retained (abstract and formal). Kohlberg (1990) argued that Piaget had a third formal stage, referred to as consolidated. Commons and Richards (1984a, 1984b) placed that stage as postformal and called it systematic. Even further, beyond Piaget's stages, Commons and associates identified three higher stages: metasytematic, paradigmatic, and cross-paradigmatic (Commons et al., 2001; Commons & Richards, 2003; M. L. Commons, personal communication, December 6, 2003).

The strengths of HCSS, the method to analyze the structure of texts that is based on MHC, are that it is designed for non-reactive analysis of texts and to be free of the domain limitations that characterized Kohlberg's method. It also overcomes the moral orientation perspective controversies raised by Gilligan, and the cultural bias generated by matching arguments to exemplars based on inputs from people who were predominantly middle class Whites (Commons et al., 2001; Dawson & Kay, 2001). Dawson and Kay, comparing HCSS scores to stages reported for 637 moral judgment interviews using Kohlberg's Standard Issue Scoring System, found 0.88 correlation. Much of the difference occurred in stages below 3. They believe that measures for those stages are strong for the structure-based HCSS but unreliable for Kohlberg's system because the sample population Kohlberg used to construct specimen responses contained no participants younger than age 10 (Dawson & Kay, 2001).

Rather than comparing responses to exemplars in a scoring manual, HCSS scores the highest order of abstraction in text elements and the logical structures coordinating the elements. This is a structural approach to analysis that can be applied to content in any domain and that attends to specifics of content only for the limited purpose of

identifying hierarchical order (Dawson & Kay, 2001). Dawson said whether a moral argument appeals to caring or justice concerns is irrelevant; the HCSS process is not sensitive to specific content. HCSS focuses solely on hierarchical complexity, a single dimension that is common to Piaget's and Kohlberg's theories, but both of them dealt with more than the hierarchical complexity dimension that HCSS focuses on. Piaget, for example, used equilibration to explain development and HCSS does not attempt such explanation (T. L. Dawson, personal communication, August 5, 2003).

While Piaget (1976) examined the child's developmental path and documented age-related progression to higher stages, the concept of progression to higher stages is complicated by the apparent flexibility of cognition in adults. Adults can operate at a stage lower than the maximum they are capable of (Commons, Krause, & Fayer, 1992; T. L. Dawson, personal communication, December 11, 2001). One reason they may do so is a belief that a particular stage will create their desired impression on a target audience (impression management) (Suedfeld et al., 2003). Thus stage of moral reasoning is a variable rather than an unwavering trait. Commons et al. (2001) asserted that performing cleaning tasks adequately requires at least Stage 7 reasoning (on Commons' scale), physical laborers' work at least Stage 8, filing clerks and typists at least Stage 9, secretaries and technicians at least Stage 10, managers, professionals, and whistleblowers at least Stage 11, and leaders, innovators, and appellate court judges at least Stage 12.

Adults also can begin performing at a stage higher than they have in the past. Commons and associates (Commons et al., 2001; Commons & Richards, 2002) reported observing five steps in transition to a higher stage: (a) perceived need to change performance (feeling of failure at current level), (b) dejection (feeling like giving up), (c) deconstruction (a search for what might be changed), (d) construction (trying changes, learning what to do and when and where to do it), (e) closure (satisfaction with success of new performance).

The HCSS scoring procedure can be used to analyze the hierarchical complexity of text whether it was created by an individual or a group, so there is no bar to using HCSS to measure pivotal leaders' vision statements at organization level. The concept of transitions is important in understanding ratings reported by HCSS scorers. "Formal" indicates performance elements rated at Stage 10. "Formal with some Systematic" indicates performance elements rated at Stage 10 plus indicators of at least one of the five transition steps leading to Stage 11.

Origin, Trend and Salient Research: Moral Reasoning Orientation

As noted above, Kohlberg's justice and rights perspective set a widely recognized standard for evaluating moral reasoning orientation. Carol Gilligan (1988a, 1998b, 1993; Gilligan & Attanucci, 1988) opened the door for research analyzing whether that perspective of justice and rights or an alternative perspective of caring and response better explained a moral reasoning event. Other work has explored age and gender differences. Some research has reported age differences—indications that younger

people may use the justice perspective more often and older people the caring orientation more often—but research generally has not documented the existence of gender differences (Colby, Kohlberg, & Kauffman, 1987; Stevens-Long & Commons, 1992). In Johnston's (1988) research, interviews of adolescents used a progressive series of probes to test whether they could switch orientations and recognize caring if they had first used justice, or vice versa. Results showed they could recognize both orientations, could switch between them, and sometimes combined them.

Application and Measures: Moral Reasoning Orientation

Following Gilligan's (1993) identification of caring as a moral reasoning orientation alternative to a justice perspective, Lyons (1982, 1988) developed a method to comparatively score justice and caring. Codable statements in interview transcript texts are extracted then categorized by orientation, either care or justice or both. The category appearing with highest frequency in a single text is scored as the predominant perspective. Since Johnston (1988) used a variation of Lyons' model to demonstrate that adolescents routinely have access to both justice and caring perspectives and may choose to use either, moral reasoning orientation may be considered a variable rather than an unwavering trait.

The answer to my question four is that two relevant streams of developmental psychology are moral reasoning stage and moral reasoning orientation.

Non-Reactive or Unobtrusive Measures

Question five asked if unobtrusive measures are useful for this research. Would more direct methods like surveys, interviews, and simulations be better?

Countering Reactivity

C. Daniel Batson (1998) opined that it was productive to study prosocial behavior by using high-impact deceptions that today will not be approved by institutional research ethics boards. He argued that alternative approaches were of limited value, including responses to hypothetical situations, self-reports, thought listing, and retrospective analysis (Batson, 1998). A weakness of such alternatives is reactivity: "Even when he is well intentioned and cooperative, the research subject's knowledge that he is participating in a scholarly search may confound the investigator's data" (Webb, Campbell, Schwartz, & Sechrest, 1966, p. 13).

Reactivity effects can be significant, witness the "Hawthorne Effect" in which investigators discovered that improved work performance may not have been caused by changes in the working conditions they thought they were studying but instead by the increased motivation workers felt because investigators were observing them (Mayo, 1945; Roethlisberger & Dickson, 1939).

Strengths and Weaknesses

Webb et al. (1966) have argued that investigators' study of archival records is

especially valuable because they have a special strength: they are a source of data untainted by interaction between investigator and research subject and therefore offer non-reactive measures. Of course, unobtrusive, non-reactive measures of real-world events do not allow control over variables available under laboratory conditions. Nonetheless, Webb et al. argued the benefits outweigh the costs. With Arnold Binder (1964), they proposed to use all available methods of research, apply judgment and intuition, and accept such data as useful even when statistical reliability fell below levels reported by empirical modes.

This study posits that content analysis of vision statements by pivotal leaders offers a true unobtrusive measure of real leader operations in the real world.

Content Analysis Facilitates Study of Catastrophe Leadership

In the study of catastrophe leadership one cannot, of course, expect verification of findings through the repetition of the event (Gould, 1989). But content analysis, as an unobtrusive and non-reactive measurement technique, permits comparison of natural events across time and cultures (Peterson, 1992; Schulman, Castellon, & Seligman, 1989; Seligman, 1998; Webb, Campbell, Schwartz, & Sechrest, 2000). The trustworthiness of findings may also be questioned where analysis is affected by investigator bias or fuzzy interpretations. For this reason, several procedural controls have been recommended, including the extraction of passages that meet criteria, the rating of passages in random order, the use of multiple raters blind to each other and to outcomes, and the measurement of intercoder reliability (Smith, 1992b; Webb et al., 1966; Winter, 2003b).

Although I refer to these techniques as "content analysis" because that has been the term commonly used in the literature (Post, 2003b; Smith, 1992b), Michael Commons used the more specific term "structure analysis." Structure examines the elements that constitute key ideas and how those elements are assembled. Moral reasoning stage is a complex example of the structure approach because it examines sub-assemblies to evaluate what hierarchical stage they achieve (personal communication, December 6, 2003; Commons et al., 2001; Commons & Richards, 2002; Dawson & Kay, 2001). The other three techniques can also be viewed as simpler structure analysis techniques (see Appendixes A – D).

The research design used here relies on a modification of previous work in this area. Using integrative complexity as a measure, Tetlock (1981) found leaders' speeches during United States presidential campaigns showed less integrative complexity than the same leaders' speeches after inauguration. The change was sharp and complexity did not further increase with time in office. In another integrative complexity study, public statements by multiple officials of nations that experienced surprise attack were examined. They showed high complexity before the attack and a decrease in complexity immediately after the attack (Suedfeld & Bluck, 1988). In a third integrative complexity study, international crisis communications of many types were evaluated, including press statements, press conference and interview responses, speeches, letters, diplomatic

cables, and diplomatic notes. Generally, results showed decline in complexity just prior to the onset of crises and rising complexity after crises began as efforts at resolution were underway (Raphael, 1982). A fourth integrative complexity study showed declines in the complexity of U.N. General Assembly speeches before outbreaks of war in the Middle East (Suedfeld et al., 1977).

So the answer to my fifth question is that unobtrusive measures fit catastrophe leadership research well. They do not raise issues of what is acceptable ethically, they do not raise issues of reactivity, and they are not limited to a moment in time and a particular place but instead can compare unrepeatable real events across time and across cultures.

Summary

In summary, leadership research usually has focused on analysis at individual level where inputs from advisors and consultants are a confounding factor. This study fits with other recent leadership studies that have focused on analysis at organization level where input from multiple persons is not seen as confounding but as part of the integrated phenomenon under study. The vision statements of pivotal leaders, built on ideas intended to represent organization members' interests and vetted by people working together to help the organization's formal leader achieve the organization's goals, do not represent the leader-as-an-individual but do represent shared leadership.

In terms of Schein's (1992) analysis of organizational culture, vision statements are an artifact that can be examined to assess two deeper levels, values and underlying assumptions.

Prosocial behavior, whether motivated by altruism, collectivism, or principism, is evidenced by a number of behaviors. This study posits that moral reasoning stage, moral reasoning orientation, explanatory style, and rumination all reflect to some degree the emergence of prosocial behavior.

Content analysis, with procedural controls, offers an opportunity to evaluate statements representing real-world leaders' reactions to real-world catastrophes using unobtrusive and non-reactive measures that permit comparison of natural events across time and cultures.

Hypotheses

Moral Reasoning Stage

HCSS is a measure of hierarchical complexity derived from developmental psychology theories of moral reasoning stages. Social psychologists have measured cognitive complexity in before-and-after studies using several methods. For example, Tetlock (1981) found leaders' speeches during United States presidential campaigns showed less integrative complexity than the same leaders' speeches after inauguration. The change was sharp, and integrative complexity did not increase further with time in office. Leaders have been shown to demonstrate less integrative complexity during stress

(Ballard, 1983; Porter & Suedfeld, 1981; Suedfeld & Tetlock, 1977; Suedfeld et al., 1977) and public statements by multiple officials of nations that experienced surprise attacks showed high integrative complexity before the attack and a decrease immediately after the attack (Suedfeld & Bluck, 1988).

Is it reasonable to expect moral reasoning stage, as measured by HCSS, to vary as integrative complexity did?

Hypothesis 1. Vision statements before a catastrophe will show complexity expressed as high moral reasoning stages. Vision statements after a catastrophe will show discrepancies in the direction of less complexity as marked by lower moral reasoning stages.

Moral Reasoning Orientation

Regardless of gender, people tend to use either the justice and rights moral reasoning orientation or the care and response moral reasoning orientation, although they can access either (Batson, 1998; Johnston, 1988; Lyons, 1982, 1983, 1988). In this study, the pivotal leaders operate at the national level, their vision statements are likely to incorporate multiple advisors' ideas (Frum, 2003; Gelderman, 1997; Ritter & Medhurst, 2003), and therefore both justice and caring orientations are likely to appear. Under conditions of stress, the influence various people exert in a group is likely to change (Janis, 1982), so changes in moral reasoning orientation in reaction to catastrophe appear probable. No data indicate direction of change in reaction to catastrophe.

Hypothesis 2. Vision statements before a catastrophe will show both justice and caring moral reasoning orientations. Vision statements after a catastrophe will show discrepancies but the direction is undefined and could be toward either justice or caring.

Explanatory Style and Rumination

In this study, election and re-election do not appear to be, and are assumed not to be, the foreground of the pivotal leaders' interest when they made the speeches selected for examination. Assuming that Franklin Delano Roosevelt modeled more negative explanatory style and more rumination as reactions to stressful environmental influences, the speeches of leaders in this study should show more negative explanatory style and more rumination after a catastrophe.

Hypothesis 3. Vision statements before a catastrophe will show some positive explanatory style. Vision statements after a catastrophe will show discrepancies in the direction of more negative explanatory style.

Hypothesis 4. Vision statements before a catastrophe will show less rumination than vision statements after a catastrophe.

CHAPTER THREE

Method

Basic Research Design

This study undertakes to compare and contrast data from 7 pivotal national leaders using rigorous latent content analysis. Vision statements are searched for changes in moral reasoning stage, moral reasoning orientation, explanatory style, and rumination.

Selection of Leaders

I selected the 7 leaders by first identifying a catastrophe—a sudden, uncontrollable event with impact so serious that a leader arguably perceived the organization or a significant part of it might not achieve fundamental goals or even survive. I selected leaders whose vision statements preceded and followed such a catastrophe.

While all of the perceptions underlying a leader's vision statement cannot truly be known, the surface content of these vision statements speaks plainly to the issue of achieving fundamental organizational goals or organizational survival. Lincoln's constant theme was preserving the Union, preserving his nation as an organization. Wilson's theme was making the world safe for democracy and his nation. Roosevelt spoke of Axis plans to divide the world and of the threat that posed to the United States. Kennedy opposed the Soviet Union's missile build-up in Cuba because it threatened global nuclear war neither nation could survive. Meir spoke of Arab nations trying to overwhelm her nation. When Argentina attacked the Falkland Islands, Thatcher spoke of fundamental goals of self-determination, and she was outraged that Argentina could believe it could get away with seizing British territory (Campbell, 2003). Thatcher biographer John Campbell suspects a secondary perception, too, that Thatcher thought loss of the islands to Argentina threatened her personal position and the survival of her government. Bush, after the September 11, 2001, terrorist attacks said, "our way of life, our very freedom, came under attack," a representation of danger to the nation and to the nation's fundamental goals (Frum, 2003, p. 126).

Identification of Vision Statements

I compiled an inventory of prepared speech texts before and after the catastrophe that are available from official or authoritative sources of good reputation for each leader. I examined these inventories to identify the subset of speeches meeting the criteria for a vision statement: accessible to substantially all organization members or their designated representatives, containing an image of the future organization, and supporting an attractive improvement over perceived alternatives. The 42 vision statements I selected are identified in Appendix I. The texts are comparable in meeting the tests to qualify as a vision statement. I avoided subjective editing by rules of type,

like whether they were statements of national direction or statements of presidential policy, or by rules of relative importance of topic, like whether they addressed global nuclear war preparedness or improvements to education, because I sought to examine each leader's pattern of actual vision statements of all types before and after a catastrophe.

Unit of Coding

In content analysis, the unit of coding varies over method. Studies of moral reasoning stage, moral reasoning orientation, and explanatory style generally use scorable extractions as the unit of coding. Studies of rumination use each sentence in the text as the unit of coding.

Comparability and Levels of Analysis

I identified comparable texts within the speeches of any one leader and between the speeches of all 7 leaders. Attention to comparability is important because differences between texts are potentially caused by differences inherent in the purpose of text rather than reactions to the catastrophe. For example, personal writings may differ from official writings, personal speaking may differ from both press conferences and formal speeches, and telephone calls and e-mail may be different too. This study focuses only on speeches, prepared in advance, and delivered by the pivotal leader. It is assumed that the texts identified are substantially accurate. While reality is that published texts, even from official or authoritative sources of good reputation, do not always exactly match as-delivered texts, published texts are an approved version identified with the leader and represent what the leader wanted to say (Winter, 2003a).

While analysis of at least four scorable extractions is "ideally required" by the explanatory style scoring manual "to generate a valid style" (Peterson et al., 1992, p. 383), Hermann (2003) has argued that adequate analysis must be based on at least 50 responses in an interview protocol. The effort here seeks to make analysis more robust by including sufficient scorable extractions to approach Hermann's standard of 50 coded units.

The research question focuses on measuring vision statements before and after a catastrophe, so one level of analysis collapses all scorable extractions before the catastrophe into one unit and all scorable extractions after the catastrophe into a second unit for analysis. Another level of analysis is available: Scorable extractions can be tallied by speech but variability in the length of speeches and the number of extractions per speech deserve attention.

Mechanically, the process for building a leader's set of vision statements begins by including the last three vision statements before the catastrophe and the first three vision statements afterward. For reference, the number of extractions from each speech will be noted.

Extractions

Only rumination scoring uses all sentences in the speech. The other three content analysis methods, when manual scoring by raters is used, first require extraction of statements that meet defined criteria. The criteria for extracting scorable statements from the speeches are as follows:

1) *Stage of moral reasoning*: For extraction a statement must contain: (a) an asserted solution to (b) what the speaker perceives as a problem, plus (c) a justification for that assertion. For scoring, extractions are further divided into propositions. After they are scored, these propositions are reassembled into statements by the investigator and evaluated to produce an overall statement score (Commons et al., 2001; Danaher & Dawson, 1999; Dawson, 1999; T. L. Dawson, personal communication, December 11, 2001).

2) *Moral reasoning orientation*: For extraction, a statement must contain: (a) a real-life moral dilemma, (b) an asserted solution, and (c) an explanation of the dilemma or an evaluation of the resolution of that dilemma (adapted from Lyons, 1982, 1988).

3) *Explanatory style*: A statement must contain a discrete occurrence that the leader perceives to have good or bad impact on the leader, the leader's in-group or organization, or the leader's country, along with an explanation by the leader—not by another person—describing what preceded the event and covaried with it from the perspective of the leader. The event and the explanation must both be good or both be bad (not a bad event with a good explanation or vice versa), and a clear causal relationship must exist between the explanation and the event (not simply sequence of events that describe without explaining and not just proof or justification of the event) (Peterson et al., 1992; Schulman et al., 1989; Seligman, 1998; Zullo, 1988; Zullo et al., 1988).

4) *Rumination*: All sentences in the text are extracted for individual scoring. The outcome is a percentage of sentences containing rumination (Zullo, 1988, 1995; Zullo & Seligman, 1990).

Reliability

Experimental Texts

For this study, the experimental texts are the leaders' vision speeches. Multiple sources acknowledge that speech texts exist in various forms (e.g., advances released to the news media before the speech, speaking texts, as-delivered texts, transcripts). Therefore, officially published versions were used on the assumption that they are substantially correct and the foundation for the leader's reputation.

Scoring

Several research assistants served as raters blind to each other and to outcomes. Raters were not told who said each passage but content clues may have allowed them to

guess. For example, who said "Yesterday, December 7, 1941—a date which will live in infamy" (Roosevelt, 1941/1950)? Some studies delete such clues, but that was impossible for this study because they are vital to the vision statements. Intercoder reliability is evaluated using the Cronbach's alpha utility of SPSS Base 10.0, release 10.0.7. Cronbach's alpha computes the ratio of average inter-item covariance to average item variance and yields estimated inter-item correlation (Cronbach, 1951; SPSS, 1999). A high alpha (approaching 1) indicates all raters appear consistent in scoring the same dimension; scores differ because raters' opinions differ, not because they are confused about how to apply rating rules. Cronbach's alpha also favors a large "sample size"; increasing the number of items tends to increase alpha. A low alpha (approaching 0 or less than 0) indicates raters appear inconsistent, as if measuring different dimensions. The higher the alpha, the higher the reliability. No general agreement on minimum acceptable alpha exists. Various sources recommend 0.68 (Satterfield, 1998), 0.7 (Yu, n.d.), 0.8 (UCLA Academic Technology Services, 2002), 0.85 (Smith, 1992b), or 0.9 (Smith, Feld, & Franz, 1992). Zullow and Seligman (1990) reported satisfactory reliability for values of 0.67 and above.

For moral reasoning stage, T. L. Dawson served as an expert rater using computer software to support HCSS scoring. For information, the scoring aid that would be used for manual rating is Appendix A. Dawson's actual method is described in Dawson and Wilson (in press).

Extractions for the other three measures were grouped by category so they could be presented to raters with instructions and scoring aids appropriate to the category (Appendixes B, C, and D). Within categories, to reduce effects of serial position, extractions were presented in random order. To reduce fatigue effects, raters were encouraged to change to scoring a different category of items every hour (Smith, 1992b; Winter, 2003b).

Data Analysis

A standard pattern of analysis for all methods has these elements: (a) number of extractions by speech, and grouped and summed in collapsed units before and after the catastrophe; (b) raters' mean scores by speech, and grouped and averaged in collapsed units before and after the catastrophe; (c) intercoder reliability; (d) paired-comparison *t*-tests; (e) graphic comparison of scores in collapsed units before and after the catastrophe; and (f) graphic comparison of scores by speech. A pilot study (Oliver, 2003) indicated too few data elements would exist to allow meaningful correlations or other rigorous studies of relationships among variables but patterns observed could be noted for future study.

A one-sample *t*-test is appropriate for testing whether the mean of a single variable is statistically different from a hypothesized population value. A paired-comparison *t*-test is appropriate for testing whether the mean of casewise differences between two variables differs from zero and often is used to evaluate before-and-after data. Paired comparison *t*-testing also returns a calculation of correlation (SPSS, 1999).

Based on these analyses, the four hypotheses can be evaluated against existing psychological theory that predicts measurable change to determine how the visions articulated by these 7 public sector leaders changed in response to these seven catastrophes. Inferences based on latent content analysis of moral reasoning stage, moral reasoning orientation, explanatory style, and rumination may provide insight into unconscious beliefs, perceptions, thoughts, and feelings that underlie vision statements leaders use to guide their organization through a human-caused catastrophe.

CHAPTER FOUR

Findings

This chapter, organized into five parts, focuses on each measure in turn: moral reasoning stage, moral reasoning orientation, explanatory style, rumination, then interactions of the four measures. Generally, findings are presented in this pattern for each measure: data collected and significant findings, possible explanations of the findings, anomalies observed, then an interim summary with hypothesis testing.

Moral Reasoning Stage

The Data Collected

Over the course of 20 months, full texts of the 42 vision statements were submitted for expert moral reasoning stage scoring to T. L. Dawson, who is supported by a computerized system for scoring hierarchical complexity (Dawson & Wilson, in press). Dawson reported the numeric scores shown in Appendix E, which ranged from 15 to 20. On Commons' (Commons et al., 2001) scale, the range is from formal to metacognitive-with-some-systematic. On Dawson's (Dawson & Wilson, in press) revised scale, the range is from consolidated abstract mappings to second transition to single principles. Figure 1 displays the scores graphically and Table 2 shows sample behaviors for stages 15 to 21.

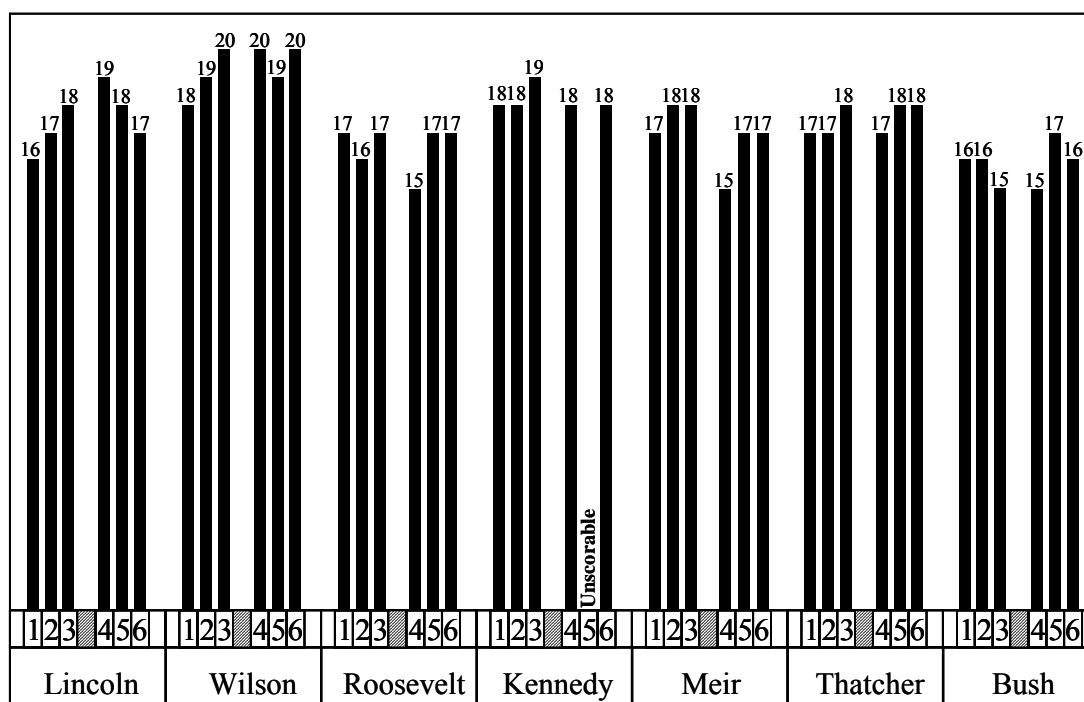


Figure 1. Stages of moral reasoning by vision, 1-3 before catastrophe and 4-6 afterward.

Table 2
Sample Behaviors for Moral Reasoning Stages 15 to 21

Dawson's Stage	Commons' Stage	Sample Behaviors
15 Consolidated Abstract Mappings	10 Formal	Linear, logical, one-dimensional: If you are attacked, then defend yourselves.
16 1st Transition to Abstract Systems	10 +	Formal with some Systematic
17 2nd Transition to Abstract Systems	10 + +	Systematic with some Formal
18 Consolidated Abstract Systems	11 Systematic	Systems words (e.g., legal system, society, the economy), relationships among variables: Organize and mobilize all the material resources of this country to supply the materials of war and serve the needs of the nation abundantly, economically and efficiently.
19 1st Transition to Single Principles	11 +	Systematic with some Metasystematic
20 2nd Transition to Single Principles	11 + +	Metasystematic with some Systematic
21 Consolidated Single Principles	12 Metasystematic	Properties of systems, systems of systems: It only remains for the two great democracies to face their task with whatever strength God may give them.

Dawson was able to score 41 of the 42 vision statements. The mean score is 17.4, the median is 17, the mode is 17, and the standard deviation (population) is 1.3. On Commons' scale, the dominant Stage 17 is systematic-with-some-formal; on Dawson's revised scale it is second transition to abstract systems. In behavior terms, it addresses systems to organize efforts but tends to rely on linear logic.

Reliability

In August 2001, when Michael Commons referred me to Theo Dawson at the University of California, Berkeley, for expert scoring of moral reasoning, her scoring rated texts using Commons' 15 stages and labels, for example, abstract, abstract with some formal, formal, formal with some systematic (T. L. Dawson, personal communication, July 15, 2002). Since then, she has made some modifications. She scored most of the texts for this study using new terms for Commons' stages, for example, consolidated abstract mappings, 1st transition to abstract systems, 2nd transition to abstract systems (see Appendix E). Step 15 is "very different" from step 16, and step 18 is "very different" from step 19. But 16/17, 17/18, 19/20, and 20/21 are more similar (T. L. Dawson, personal communications, November 3, 2003; November 17, 2003). These changes signal an evolution in her scoring of moral reasoning but I had to freeze the scoring method at some point. I chose August 2003 because that was when she completed scoring the final set of vision statements for this study and provided a key analytical parameter: "Differences between scores of more than 1/3 of a complexity order are statistically meaningful" (T. L. Dawson, personal communication, August 5, 2003).

Dawson and Wilson (in press) reported the reliability of the computerized scoring system, as measured by a Rasch analysis that is equivalent to Cronbach's alpha, ranges from 80% to 97% within one-half complexity order and from 98% to 100% within one full complexity order. On the Dawson scale, complexity orders appear every third integer: 15 = abstract mappings, 18 = abstract systems, 21 = single principles. The intermediate numbers, for example 16 and 17, 19 and 20, represent one-third transitions to the next stage.

I adopted the one-half complexity order criterion supported by a Rasch analysis equivalent to 80% to 97%. *t*-tests using my data, which are reported later, supported my treatment of adjacent numbers, for example 15 and 16, as similar scores with no significant difference for my data (although they are one-third of a complexity order apart) they are within one-half complexity order and within the Rasch 80% to 97% reliability range reported above for this rating system. My treatment of non-adjacent numbers ($\Delta > 1$) as significantly different scores for my data (when they are more than one-half complexity order apart) appeared helpful. Since the freeze date, Dawson has told me that most of their research showed "scores to be reliable within one-third of a complexity order" but she had read my research and conclusions and judged that the new information did not affect them (T. L. Dawson, personal communication, November 3, 2003).

Inspection of Figure 1 suggests pattern effects: variations of one step between adjacent vision statements look common, perhaps normal. Clearly these vision statements did access different stages. Lincoln showed wide variation, from 16 to 19, but adjacent visions scored only one step apart. Wilson showed smaller variation, from 18 to 20, again with adjacent visions not more than one step apart. The one-step pattern was

true also for Kennedy and Thatcher. The three anomalies are Roosevelt, whose fourth vision—his speech to Congress after the Pearl Harbor attack—showed a two-step drop followed by a two-step rise; Meir, whose fourth vision—her speech announcing an invasion—showed a three-step drop followed by a two-step rise; and Bush, whose fourth vision matched his third vision's low score, 15, and who then showed a two-step increase on vision five. As detailed below, this visual test is confirmed statistically if the standard is one-half order. If the standard were one-third order, then 22 one-step changes (64.7%) would be significant.

I tested the 41 scores using a one-sample *t*-test to determine if the mean differs from a hypothetical population value. The test showed a *p* value of < 0.0005 , indicating that the null hypothesis is rejected at the .01 level. The mean is 17.3902. The confidence interval for this mean extends from 16.8325 to 17.9480, which rounds to the integers 17 and 18, supporting my conservative assessment that adjacent numbers may be similar scores but non-adjacent numbers ($\Delta > 1$) should be seen as distinct scores when they exceed one-half stage.

Statistically Significant Findings

Previous research created a reasonable expectation that moral reasoning stage, a measure of complexity, would decrease after a catastrophe. Research using a different measure of complexity, integrative complexity, showed that leaders' complexity scores decreased during stress (Ballard, 1983; Porter & Suedfeld, 1981; Suedfeld & Tetlock, 1977; Suedfeld et al., 1977) and that public statements by multiple officials of nations that experienced surprise attacks showed a decrease in complexity immediately after the attack (Suedfeld & Bluck, 1988). Reasonable theories explain why moral reasoning stage would decrease after a catastrophe: crisis situations, like being attacked, often require only the abstract stage (Dawson Stage 12) for response; performance decreases because working memory decreases as emotion increases (M. L. Commons, personal communications, June 9 and December 22, 2003).

Expert scoring of these 7 pivotal leaders' vision statements showed three outcomes that appear significant. The first may be seen as a norm for this group: Five of the 7 leaders (71.4%) showed no significant change immediately after a catastrophe. However, 2 leaders did show the expected decrease in moral reasoning stage immediately after a catastrophe, and 3 leaders showed low moral reasoning stage immediately after a catastrophe followed by a significant increase in stage in their second vision statement after the catastrophe, which also is similar to a pattern seen in previous research.

The norm is set by Lincoln, Wilson, Kennedy, Thatcher, and Bush. For these five leaders (71.4%), the moral reasoning stage of the vision statement immediately following the catastrophe is within one integer ($\Delta \leq 1$) of the statement immediately preceding the catastrophe.

The first deviation from the norm appears in data for Roosevelt, where the post-catastrophe vision statement scored two integers below the pre-catastrophe vision ($\Delta =$

2), and for Meir, where the post-catastrophe vision statement scored three integers below the pre-catastrophe vision ($\Delta = 3$).

The second deviation from the norm appears in data for Roosevelt, for Meir, and for Bush. For all 3, the first post-catastrophe vision statement scored low and was followed by another vision statement scored two integers higher ($\Delta = 2$). For all 3, the change was from Stage 15 to Stage 17. This deviation is consistent with findings reported by Raphael (1982), who saw complexity scores decline just prior to the onset of crises and then rise subsequently, which Raphael interpreted as related to crisis resolution efforts.

Several hypotheses can be advanced to explain these results. For the 5-leader norm, evidence indicates that the moral reasoning stage scores basically fit a normal curve. Figure 2 shows the frequency distribution for all 41 scores. The pattern approximates a normal curve, skewed slightly toward the lower stages. The skew is confirmed mathematically by the median at 17 being slightly lower than the mean of 17.4.

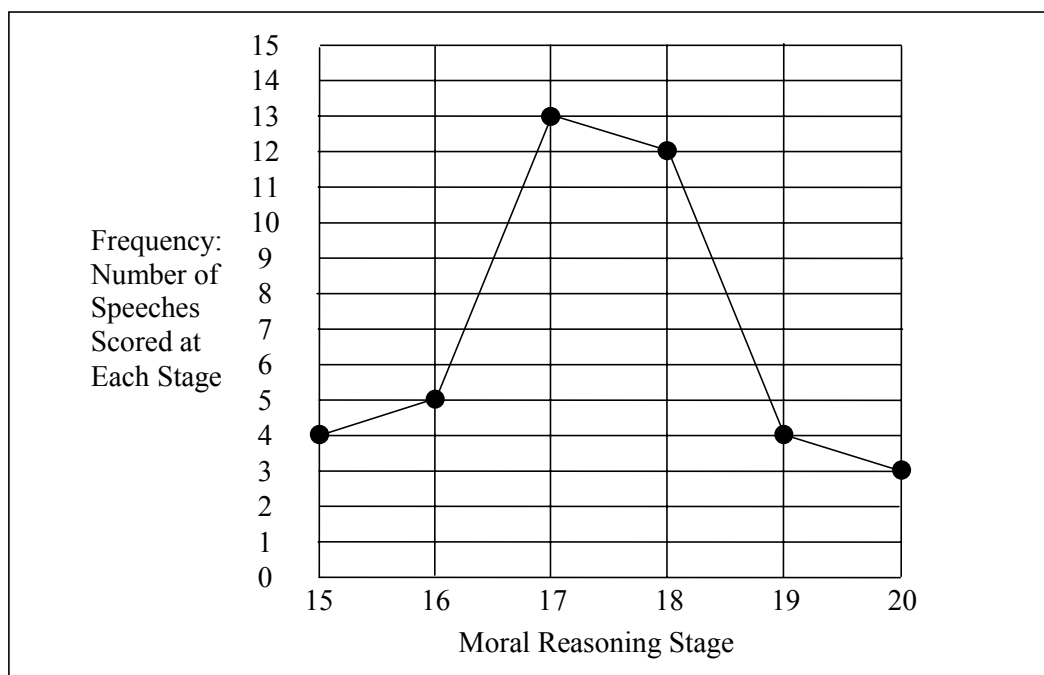


Figure 2. Frequency distribution of moral reasoning stages in the 41 scored visions.

If observed in a randomly selected population, many types of data approximate a normal curve within an interval determined by the extremes of the population examined (e.g., height, weight, annual income, aptitude test scores). Such data are marked by independent events. That moral reasoning stage scores for these 41 visions approximate

a normal curve implies they can be viewed as independent events. Observe also that removing the two significant decreases, Roosevelt's and Meir's fourth vision statements, removes 2 scores of 15 from the distribution, reduces the skew, and improves the remaining scores' approximation of the normal curve. One explanation for appearance of a normal curve in these scores might be that, with respect to moral reasoning stage, these leaders represent a sample selected randomly from the population, that is moral reasoning stage was not a standard criterion they had to meet to qualify as a leader.

The normal curve similarly appears when the frequency distributions for vision statements made before a catastrophe are plotted separately from those for vision statements made afterward (Figure 3). If Roosevelt's and Meir's fourth vision statements, with scores of 15, are removed from the post-catastrophe distribution, the upward spike at the low extreme will disappear so both curves more resemble a normal distribution.

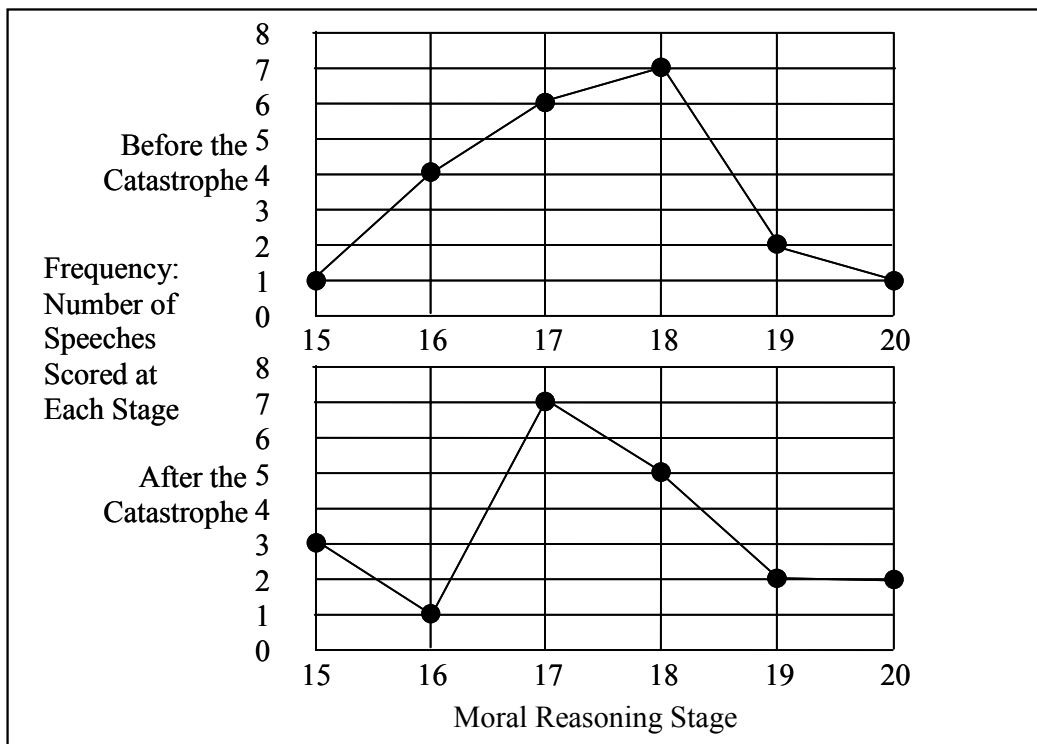


Figure 3. Stages of moral reasoning by vision before and after a catastrophe.

The first deviation from the norm, the decreases in moral reasoning stage immediately after a catastrophe for Roosevelt and Meir ($\Delta > 1$), are interesting and consistent with expectations based on previous research. The decreases may relate to qualitative differences in the catastrophes these two leaders faced. Both Roosevelt and Meir reacted to foreign nations' overt military attacks on their homeland (Meir, 1973a; Roosevelt, 1941/1950). None of the other leaders faced exactly that issue. Lincoln foresaw the attack on Fort Sumter by insurgents internal to his nation (Lincoln,

1861/1953), Wilson resisted believing that Germany was conducting unrestricted submarine warfare that affected the United States (Wilson, 1917/1983), Kennedy detected a build-up of Soviet offensive weapons in Cuba before they became operational (Kennedy, 1962a), Thatcher responded to an attack by Argentina on British islands thousands of miles distant from her homeland (Thatcher, 1982a), and Bush faced attacks by terrorists whose sponsoring organization was still to be identified at the time he spoke (Bush, 2001f).

The second deviation from the norm, the increases in moral reasoning stage in the second vision after a catastrophe for Roosevelt, Meir, and Bush from 15 to 17 may indicate deliberate framing of the speech immediately following a catastrophe to describe the situation in simple terms, for clarity or impression management, followed by another vision that elaborates on implications and complexities. Some evidence supporting this view appears in an unsolicited comment that expert scorer T. L. Dawson offered on Bush's vision statement four: "This is a carefully crafted speech [Stage 15], written by professional speech writers to be comprehensible to anyone over the age of 14 or 15. It has clearly been written by someone who could write at the systematic order [Stage 18], but has chosen not to do so" (personal communication, December 10, 2001).

Another explanation is that the stage scoring might be inaccurate. Two factors support the accuracy of the scoring. One is the testing that demonstrated the computerized scoring system is reliable (Dawson & Wilson, in press). Another is a pilot study I conducted in December 2002 using human raters to score a subset of the current data using a manual, two-stage scoring protocol. Those raters reported scores for Thatcher and Bush similar to the current expert scoring (Oliver, 2003).

Why does the 5-leader norm not show the decrease in stage consistent with expectations based on previous research? That it does not is qualitatively significant in the face of evidence that stage has decreased for individuals in the past and that reasonable justifications for stage decrease exist. Is it possible that shared leadership, with its implied influence by multiple people, might influence the vision statement in the direction of displaying an "even keel"? No direct test for this hypothesis is available in the data; triangulation from other information must be sought.

Other analyses showed no significant patterns. For the entire group of 7 leaders, significance testing using a paired-comparison *t*-test to evaluate the means before and after a catastrophe showed a *p* value of 0.863, indicating the null hypothesis is accepted at the .05 level. Significance testing using a paired-comparison *t*-test to evaluate just these 7 leaders' one vision statement immediately preceding the catastrophe and the one vision statement immediately following it showed a *p* value of 0.143, indicating that the null hypothesis is accepted at the .05 level. Collapsing the data to consolidate the pre-catastrophe vision statements into one group and the post-catastrophe vision statements into another group yielded no significant finding (Figure 4). Four leaders' moral reasoning stages appeared to increase slightly after the catastrophe (Lincoln, Wilson, Thatcher, and Bush) and 3 leaders' moral reasoning stages appeared to decrease slightly (Roosevelt, Kennedy, and Meir). Only Meir showed a change that appeared to be greater

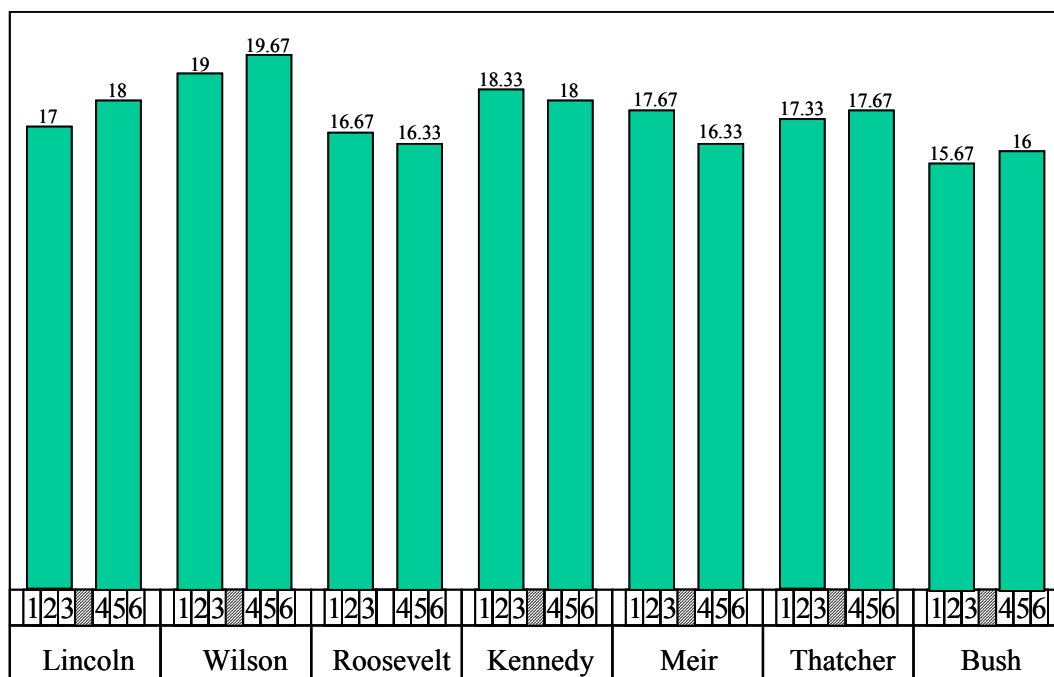


Figure 4. Stages of moral reasoning for vision statements aggregated by period: pre- or post-catastrophe.

than one integer ($\Delta = 1.34$) and even that change, rounded to the nearest integer, has a value of 1. Paired-comparison t -tests showed none of the variations is significant. Overall, scores for the group of 7 leaders showed a p value of 0.879. Scores for Meir showed a p value of 0.057. Both indicate the null hypothesis is accepted at the .05 level.

Following up on previous analysis that indicated the data fit a normal curve, Figure 5 shows that the frequency distribution for collapsed scores tallied by interval (e.g., 15 to < 16, 16 to < 17) suggests a normal curve but lacks the detail afforded by the 41 scores used for Figure 2.

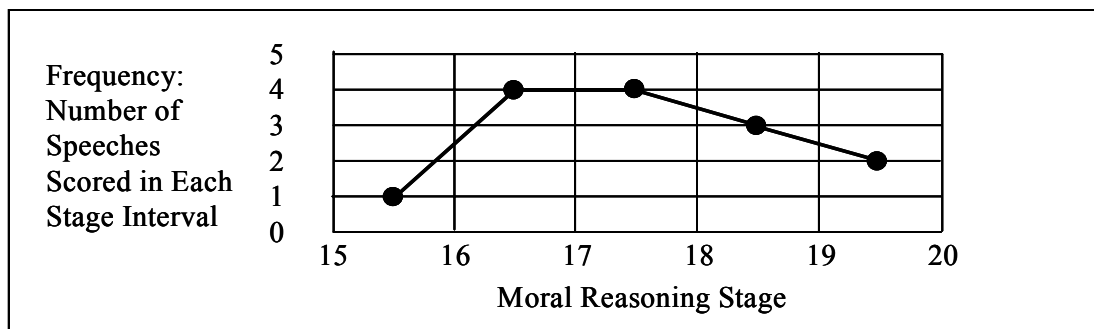


Figure 5. Frequency distribution of moral reasoning stages for vision statements aggregated by period: pre- or post-catastrophe.

Anomalies

One readily visible anomaly is Kennedy vision five because it met this study's criteria to be a vision statement but contained no text meeting the criteria to be scorable for moral development stage. Its vision described future international inspection of Cuba-bound cargo and restoration of peace in the Caribbean but did not present reasoning that explained how those vision elements would be achieved (Kennedy, 1962b). To be scored for moral reasoning stage, a text must present a problem requiring reasoning, an asserted solution, and a rationale for that solution, all elements that are not vision statement criteria. Only 245 words long, vision five followed close on the heels of vision four and was an update announcing that the Soviets were removing offensive weapons from Cuba. The speaker possibly viewed it as a continuation of vision four and built on reasoning already presented there.

Are the significant outcomes robust, or are they possibly distorted by extremes? Some visions are longer. Does a long vision overwhelm short visions? For all four measures, the process is conservative. It does not give more weight to a long vision or a vision with more extractions. Each vision is rated equally; each vision has a weight of one. For moral reasoning stage, a test of the influence that long visions have can be conducted using a paired-comparison *t*-test to determine if the mean of casewise differences between long visions and short visions differs from zero. The median number of sentences in the 41 scorable visions is 102.5. Of the 41 scorable vision statements, 20 have fewer than 102.5 sentences (designated the low set) and 21 have more (designated the high set). The paired-comparison *t*-test for that split shows a *p* of 0.860, indicating the null hypothesis is accepted at the .05 level. I conclude that the evidence shows the long visions do not change the significant results for moral reasoning stage.

Summary and Hypothesis Testing

Generally, moral reasoning stage scores for this group of 7 leaders appeared to approximate a normal curve. The norm set by 5 leaders was no significant change in moral stage following a catastrophe. Significant differences were found for Roosevelt and Meir, whose post-catastrophe vision statements addressed foreign nations' overt attacks on their homeland. Significant differences were found for Roosevelt, Meir and Bush in the second speech following a catastrophe, possibly caused by deliberate simplification of the first speech followed by elaboration on implications and complexities in the second.

Hypothesis 1 was that vision statements before a catastrophe will show complexity expressed as high moral reasoning stages and vision statements after catastrophe will show discrepancies in the direction of less complexity as marked by lower moral reasoning stages.

Analysis of the vision statements used for this study did not support the hypothesis. This does not refute previous research findings that created the expectation

expressed in the hypothesis. Those studies measured texts selected in a different way than the vision statements used here, may have examined texts attributable to an individual instead of to shared leadership, used measures other than the hierarchical complexity used here, and used scoring techniques different than those used by T. L. Dawson. Michael Commons described hierarchical complexity as more precise and integrative than previous measures used by Piaget and Kohlberg, and as more precise than cognitive complexity and integrative complexity previously used by other researchers (personal communication, June 9, 2003).

Moral Reasoning Orientation

The Data Collected

As part of a 1-day session, 115 scorable extractions from the 42 vision statements were submitted in random order to a panel of nine raters blind to each others' scoring and to possible outcomes of the study. Each extraction was scored by three raters on a 4-point scale: only caring, mostly caring, mostly justice, only justice. An example from training materials may clarify the differences (adapted from Johnston, 1988).

- Situation: It was growing cold, and a porcupine was looking for a home.
- Conflict: He found a most desirable cave but saw it was occupied by a family of moles. The cave was small and every time the moles moved around they were scratched by the porcupine's sharp quills. "Pray leave," they said to the porcupine. "Oh no," said the porcupine. "This place suits me very well."
- Solution 1: The porcupine definitely has to go. It's the moles' house.
- Solution 2: Both the porcupine and the moles should try to get together and make the hole bigger.

In training, the situation and conflict are presented once with solution 1 and later with solution 2. Solution 1, focused on rights and rules, makes it an "only justice" item; solution 2, focused on relationships and caring, makes it an "only caring" item. Were a third solution created, such as, "The porcupine and the moles should try harder to get along but if they can't then the porcupine should move out," then the item might be rated either caring with some justice or justice with some caring.

For statistics, "only caring" and "only justice" responses were assigned twice the weight of "mostly caring" and "mostly justice." The mean of the three ratings was used as the extraction's overall score. The extractions for each vision statement were recombined and the consolidated mean became the vision statement's score. Table 3 shows the number of extractions for each vision statement. Of the 42 vision statements, 40 were scorable and two had no content meeting extraction criteria. For the 40 scorable vision statements, the number of extractions ranged from 1 to 9, with a mean of 2.88, median of 2, mode of 1, and a standard deviation (population) of 2.24.

Table 3

Number of Moral Reasoning Orientation Extractions for Each Vision Statement

Before the Catastrophe		After the Catastrophe	
Text	No. of Extractions	Text	No. of Extractions
Lincoln 1	1	Lincoln 4	9
Lincoln 2	3	Lincoln 5	7
Lincoln 3	7	Lincoln 6	2
Wilson 1	1	Wilson 4	4
Wilson 2	1	Wilson 5	1
Wilson 3	2	Wilson 6	1
Roosevelt 1	2	Roosevelt 4	1
Roosevelt 2	1	Roosevelt 5	3
Roosevelt 3	2	Roosevelt 6	1
Kennedy 1	3	Kennedy 4	1
Kennedy 2	2	Kennedy 5	0
Kennedy 3	2	Kennedy 6	4
Meir 1	2	Meir 4	0
Meir 2	2	Meir 5	1
Meir 3	1	Meir 6	3
Thatcher 1	14	Thatcher 4	5
Thatcher 2	1	Thatcher 5	4
Thatcher 3	2	Thatcher 6	4
Bush 1	1	Bush 4	2
Bush 2	2	Bush 5	7
Bush 3	2	Bush 6	1

Numeric scores (see Appendix F) ranged from 100% caring to 100% justice orientation. Figure 6 displays the polar scores graphically. Polar scores apply a "predominance of mode" rule (Lyons, 1988, p. 38); if a majority of extracts are scored justice, then the recombined speech is scored justice. On two vision statements, raters reported an exact balance: half of the extracts scored as caring and half as justice.

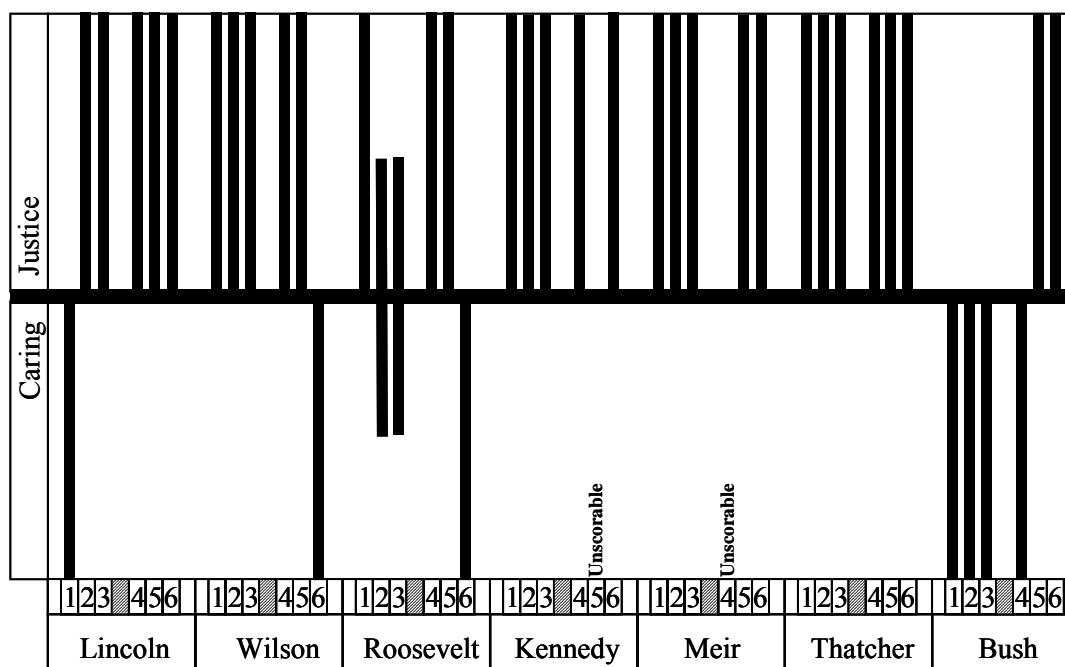


Figure 6. Moral reasoning orientation polar scores for the 42 vision statements.

Reliability

Cronbach's alpha is only 0.4092, which indicates these raters were only moderately consistent in interpreting how these extracts fit the scoring criteria. Significance testing of the 40 scored visions using one-sample *t*-tests to separately determine if the means for caring and justice differ from a hypothetical population value showed a *p* value of < 0.0005 in both cases, indicating that the null hypothesis is rejected at the .01 level. The mean for caring is 0.3453 with a 95% confidence interval from 0.2376 to 0.4531, an interval of 0.2155. The mean for justice is 1.0252 with a 95% confidence interval from 0.8676 to 1.1828, an interval of 0.3152.

Significant Findings

Previous research created a reasonable expectation that vision statements would include both extractions scored caring and extractions scored justice. Research has shown that people tend to prefer either the care and response moral reasoning orientation or the justice and rights orientation, although they can access either and use both (Batson, 1998; Gilligan, 1988a; Johnston, 1988; Lyons, 1982, 1983, 1988). The national-leader vision statements examined in this study appear likely to incorporate multiple advisors' ideas (Frum, 2003, Gelderman, 1997; Ritter & Medhurst, 2003), so both caring and justice orientations are likely to appear. Under conditions of stress, the influence various people exert in a group is likely to change (Janis, 1982), so the hypothesis forecast change after a catastrophe but the direction was undefined.

Raters' scoring of these 7 pivotal leaders' vision statements showed three outcomes that appear significant. The first can be seen as a norm: 5 leaders (71.4%) displayed an increase in justice orientation after a catastrophe. The second outcome is a deviation from that norm: 1 leader showed change in the opposite direction, an increase in caring orientation. The total is 6 of the 7 leaders (85.7%) showed a post-catastrophe discrepancy, with 5 changed toward justice and 1 toward caring. The third outcome is a significant anomaly: Bush's dramatic change from caring to justice orientation.

The norm is set by Lincoln, Wilson, Roosevelt, Kennedy, and Bush. For these 5 leaders (71.4%), the justice orientation of the vision immediately following the catastrophe increased. Significance testing of that change using a paired-comparison *t*-test showed a *p* value of 0.044, indicating the null hypothesis is rejected at the .05 level.

The deviation from the norm is set by Thatcher, whose caring orientation increased by 0.5144, significant because the change is greater than the 95% confidence interval ($\Delta > 0.2155$).

I tested the anomaly, Bush's change from caring to justice orientation, using paired-comparison *t*-tests to compare Bush's pre-catastrophe speeches to his post-catastrophe speeches. The change in justice scores was significant at the .05 level (*p* = 0.024). Figure 7 supports the statistical finding by showing that five of Bush's six vision statements contained strong caring elements but only the post-catastrophe speeches contained strong justice elements.

One hypothesis to explain the norm and the deviation is that the changes indicate deliberate framing to garner followers' support. No direct test for this hypothesis is available in the data; triangulation from other information must be sought.

Another hypothesis is that the orientation scoring might be inaccurate. One factor supporting that hypothesis is the moderate level of Cronbach's alpha. While 20% of the visions received ratings at both extremes of the scale, only caring and only justice, which would seem to be mutually exclusive categories, that is not evidence of scoring inaccuracy for two reasons. First, visions include multiple extractions and each extraction is scored independently. Second, people use both orientations, so it is reasonable to find both orientations in the same vision. One factor supporting the accuracy of the scoring is a pilot study I conducted in December 2002 using a different panel of raters to score a subset of the current data using the same manual, two-stage scoring protocol. Those raters reported scores for Thatcher and Bush similar to the current panel's scores (Oliver, 2003).

The Bush anomaly possibly is explained as deliberate framing to garner followers' support. The first three caring vision statements are part of a series of brief, pre-recorded Saturday morning radio addresses to the nation, each with a single theme. These advocated social welfare and education reforms. The fourth caring statement, immediately following the terrorist attacks, praised responses to the emergency. Frum (2003) attributed the fourth statement principally to Bush's senior communication advisor, Karen Hughes, saying words others had drafted were "chucked and replaced by Karen Hughes" (p. 126).

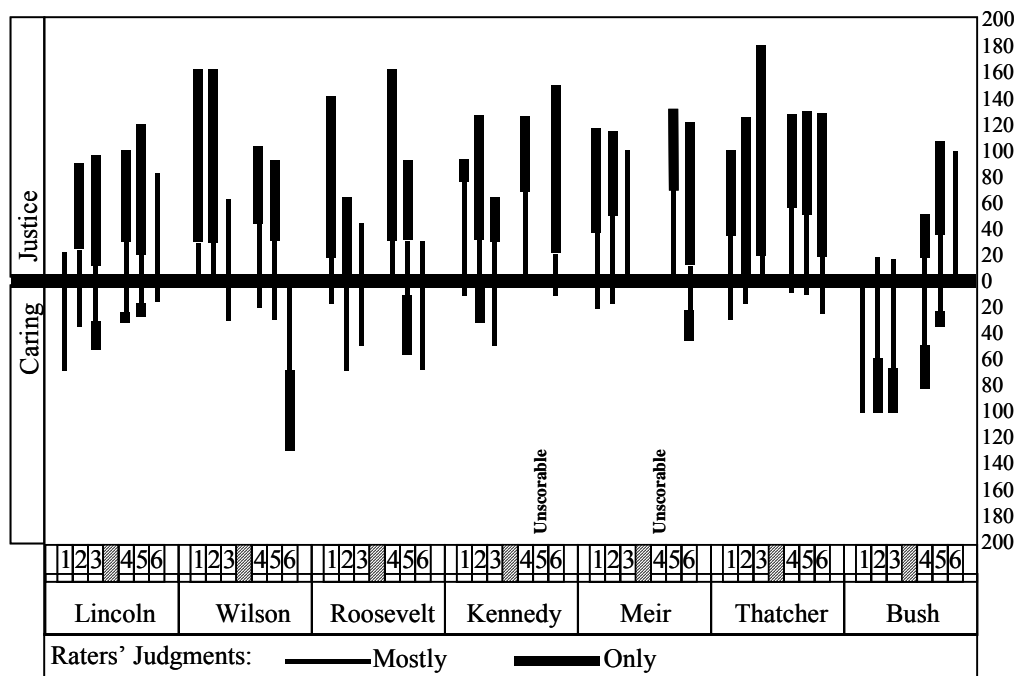


Figure 7. Moral reasoning orientation ratings of 42 vision statements on 4-point scale.

Frum (2003) criticized the caring, compassionate tone of the speech and wished it had been justice oriented, a war speech: "Americans had been slaughtered in a horrible fiery massacre. The country was stunned, terrified, furious, and ready for war against anyone and everyone in any way connected to the attack" (p. 128). Frum reported that such criticism caused Hughes to back away from drafting vision five, Bush's September 20 address to a joint session of Congress and to the nation, the first of his justice-oriented visions. Michael Gerson chaired the drafting of that speech. The change in moral orientation did garner support, according to Frum. A poll after September 11 showed that only about half the country felt Bush could cope with the crisis but a poll after the September 20 address showed "levels of trust no leader in American history had ever previously achieved" (p. 148). The sixth speech, following 2 days later, focused on overcoming the nation's economic plunge but Frum did not report who influenced it.

Other analyses showed no significant patterns. Figures 8 and 9 graphically display the results of collapsing data into two groups that consolidate the pre-catastrophe vision statements and the post-catastrophe visions. Figure 8 shows that overall ratings of collapsed vision statements are justice orientation, with only one exception: Bush before the catastrophe. Figure 9 shows that Bush's pre-catastrophe visions were strongly caring and that raters saw "only justice" orientation far more often than they saw "only caring" orientation.

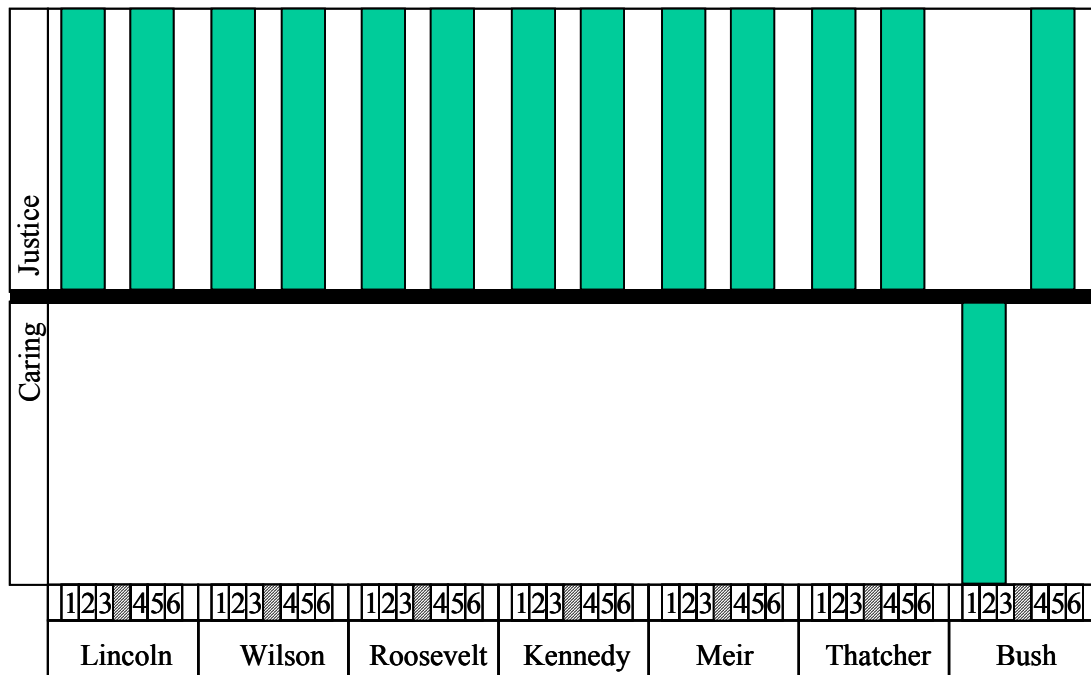


Figure 8. Moral reasoning orientation scores (polar) for vision statements aggregated by period: pre- or post-catastrophe.

Anomalies

Kennedy's fifth speech and Meir's fourth speech were unscorable because they contained no text that met the criteria for extraction: Present a real-life moral dilemma, plus an asserted solution, plus an explanation of the dilemma or an evaluation of the resolution of that dilemma. As noted before, Kennedy's fifth vision was brief, actually an extension of his fourth vision. It presented no moral dilemma. Meir's fourth vision reported that Egypt and Syria attacked Israel today and presented no explanation or resolution of the dilemma (Meir, 1973a); perhaps the attack seemed irrational, inexplicable, and unresolvable at the moment.

While it is tempting to focus on Roosevelt's second and third vision statements for scoring equally caring and justice, Figures 7 and 9 show the norm for these American presidents was to show considerable caring. Lincoln, Wilson, Roosevelt, Kennedy, and Bush all showed frequent caring. The deviations are scores of the two non-American leaders, Meir and Thatcher. Both had high justice orientation. One explanation may be that the presidents used a "presentation" style but Meir and Thatcher were both prime ministers speaking primarily in a parliamentary forum, perhaps using a discussion or debating style. Another explanation might relate to image setting. White House staff members expressed some negative reaction to Bush's September 11, 2001, address to the nation, a vision statement scored as having caring moral reasoning orientation. Critics on

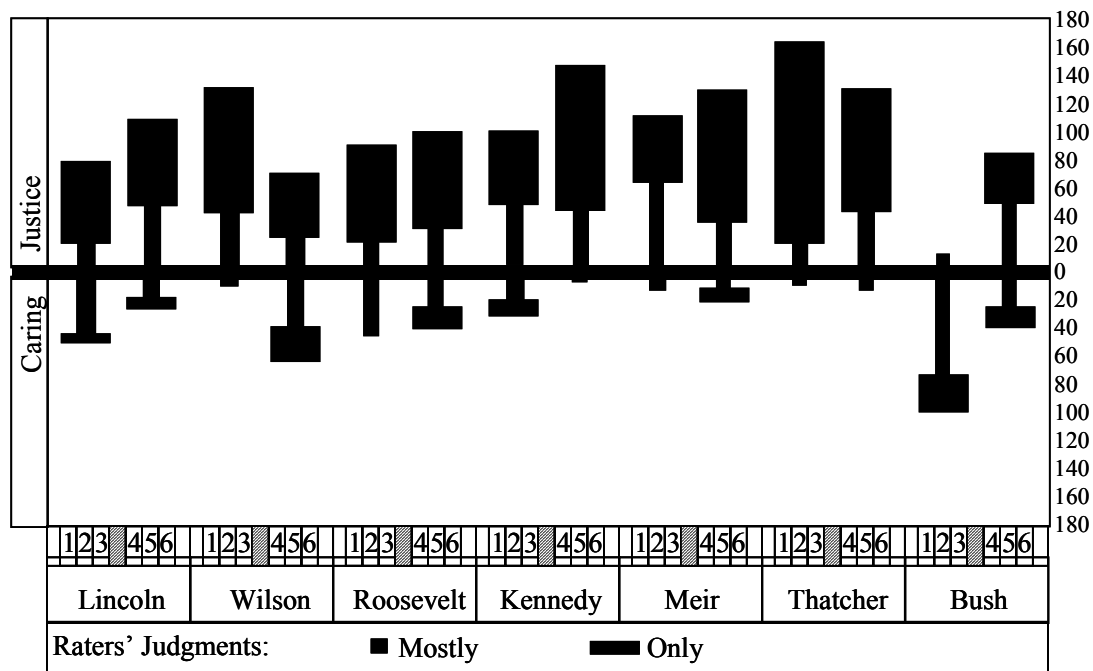


Figure 9. Moral reasoning orientation ratings on 4-point scale for vision statements aggregated by period.

the White House staff, perhaps influential members of the shared leadership group, said that speech presented an image of Bush as vulnerable and not in command (Frum, 2003). Bush switched from caring to justice orientation; perhaps Meir and Thatcher chose justice orientation for the same reason: belief that it better created an image of invulnerability and powerful command. Such belief could be based on one or more of a number of factors, such as political experience, antipathy toward gender stereotypes, or advice from consultants. Indeed, before the Falkland Islands were attacked, Thatcher already had "taught the public to see her as the Iron Lady . . . a resolute defender of British interests and British pride" (Campbell, 2003, p. 126).

Are the significant outcomes robust, or are they possibly distorted by extremes? As noted above, because some visions have more scorable extractions a question is raised whether they might overwhelm visions with three or four extractions. For all four measures, the conservative process does not give more weight to a long vision or a vision with more extractions. Each vision is rated equally; each vision has a weight of 1. For moral reasoning orientation, a test of the influence that visions with many extractions have can be conducted using a paired-comparison *t*-test to determine if the mean of casewise differences between visions with few extractions and those with many extractions differs from zero.

The median number of moral reasoning orientation extractions is two (12 at that number). Of the 40 scorable vision statements, 13 have fewer than two extractions and 15 have more than two extractions. If the 12 with two extractions are split, so the first 7

are added to the low-extractions set and the last 5 are added to the high-extractions set, then the mean of casewise differences between the low-extractions set and the high-extractions set can be tested. The paired-comparison *t*-test for that split shows a *p* of 0.873, indicating the null hypothesis is accepted at the .05 level. I conclude that the evidence shows the visions with many scorable extracts do not change the significant results for moral reasoning orientation.

Summary and Hypothesis Testing

Generally, the leaders showed an increase in justice orientation following a catastrophe but 1 leader showed an opposite change, toward caring. Bush dramatically changed from caring to justice orientation and the change occurred when presidential advisor Hughes withdrew from drafting his speeches. Generally, the norm for these American presidents was to show considerable caring. The 2 non-American leaders, Meir and Thatcher, showed higher justice orientation. One possible explanation is that they speak in a different forum, a parliamentary forum. Other possible explanations include gender pressures or image reasons.

Hypothesis 2 was that vision statements before a catastrophe will show both justice and caring moral reasoning orientations. Vision statements after a catastrophe will show discrepancies but the direction is undefined and could be toward either justice or caring. Analysis of the vision statements used for this study supports the hypothesis. Six of the 7 leaders (85.7%) showed a post-catastrophe discrepancy, with 5 changing toward justice and 1 toward caring.

Explanatory Style

The Data Collected

As part of a 1-day scoring session, 423 extractions from the 42 vision statements were submitted in random order to a panel of nine raters blind to each others' scoring and to possible outcomes of the study. Each extraction was scheduled to be scored by three raters and two-thirds were. One-third received only two scores due to a medical problem that curtailed one rater's participation.

Explanatory style scoring tasks raters to score each extraction three times: once for internality versus externality, once for stability versus instability, and once for globality versus specificity. Each characteristic is scored on a scale from 1 to 7; the midpoint, 4, is considered neutral. Scores below 4 favor one polar quality and scores above 4 favor the other. Each characteristic can be charted and all three scales can be combined by summing the scores. When combined, the scores can range from 3 to 21; the midpoint, 12, is considered neutral. Scores below 12 favor one polar quality and scores above favor the other. A complication is that explanatory style requires occurrences the speaker perceives as having bad impact to be processed separately from occurrences the speaker perceives as having good impact. The polar qualities for good occurrences are opposite the polar qualities for bad occurrences. For example, a score of

21 is highly negative for a bad occurrence but highly positive for a good occurrence.

Table 4 shows the number of extractions for each vision statement. Of the 42 vision statements, 36 were scorable and six had no content meeting extraction criteria. For the 36 scorable vision statements, the number of extractions ranged from 1 to 40, with a mean of 11.75, median of 10, modes at 1, 2, and 10, and a standard deviation (population) of 9.78.

For numeric scores, see Appendix G. Table 5 summarizes the score data. Overall, 144 good extractions and 279 bad extractions were scored. The good extractions tended to score slightly above the scale's mid-point, which for good events is in the positive direction. The bad extractions tended to score below the scale's mid-point, which for bad events is the positive direction. The exception is the globality versus specificity scale, which showed the reverse pattern.

Figures 10-13 display the scores graphically. In all of these charts, columns rising above the mid-point indicate positive direction and columns descending below the mid-point indicate negative direction.

Table 4

Number of Explanatory Style Extractions for Each Vision Statement

Before the Catastrophe		After the Catastrophe	
Text	No. of Extractions	Text	No. of Extractions
Lincoln 1	10	Lincoln 4	40
Lincoln 2	7	Lincoln 5	35
Lincoln 3	19	Lincoln 6	1
Wilson 1	2	Wilson 4	0
Wilson 2	5	Wilson 5	1
Wilson 3	29	Wilson 6	19
Roosevelt 1	25	Roosevelt 4	23
Roosevelt 2	1	Roosevelt 5	0
Roosevelt 3	7	Roosevelt 6	0
Kennedy 1	13	Kennedy 4	17
Kennedy 2	15	Kennedy 5	1
Kennedy 3	9	Kennedy 6	16
Meir 1	12	Meir 4	15
Meir 2	10	Meir 5	0
Meir 3	0	Meir 6	2
Thatcher 1	24	Thatcher 4	11
Thatcher 2	10	Thatcher 5	0
Thatcher 3	4	Thatcher 6	3
Bush 1	10	Bush 4	5
Bush 2	2	Bush 5	13
Bush 3	5	Bush 6	2

Table 5
Summary of Explanatory Style Ratings

Characteristic	Mean	Median	Std Dev (Pop)
Internality/Externality (Bad)	2.81	2.50	1.42
Stability/Instability (Bad)	3.63	3.67	1.08
Globality/Specificity (Bad)	4.26	4.33	1.41
Combined (Bad)	10.70	11.00	2.36
Internality/Externality (Good)	4.25	4.00	1.50
Stability/Instability (Good)	4.40	4.33	1.24
Globality/Specificity (Good)	3.94	4.00	1.39
Combined (Good)	12.58	12.67	2.67

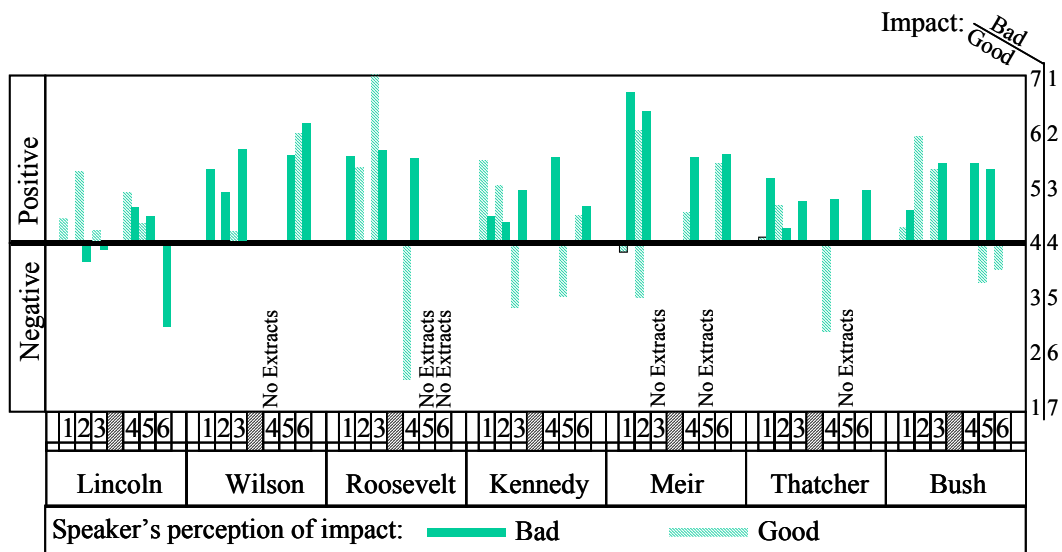


Figure 10. Explanatory style internality versus externality component for 42 vision statements.

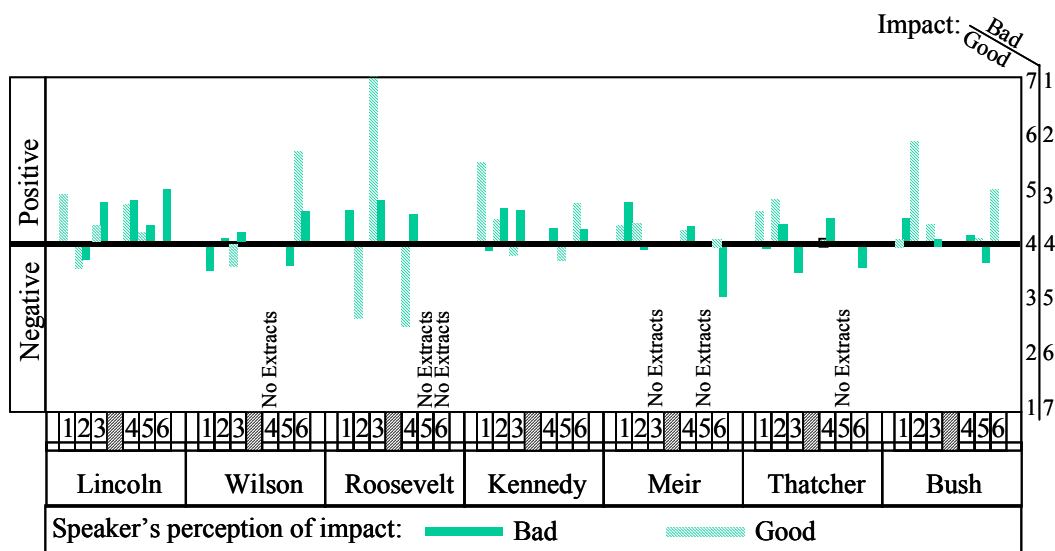


Figure 11. Explanatory style stability versus instability component for 42 vision statements.

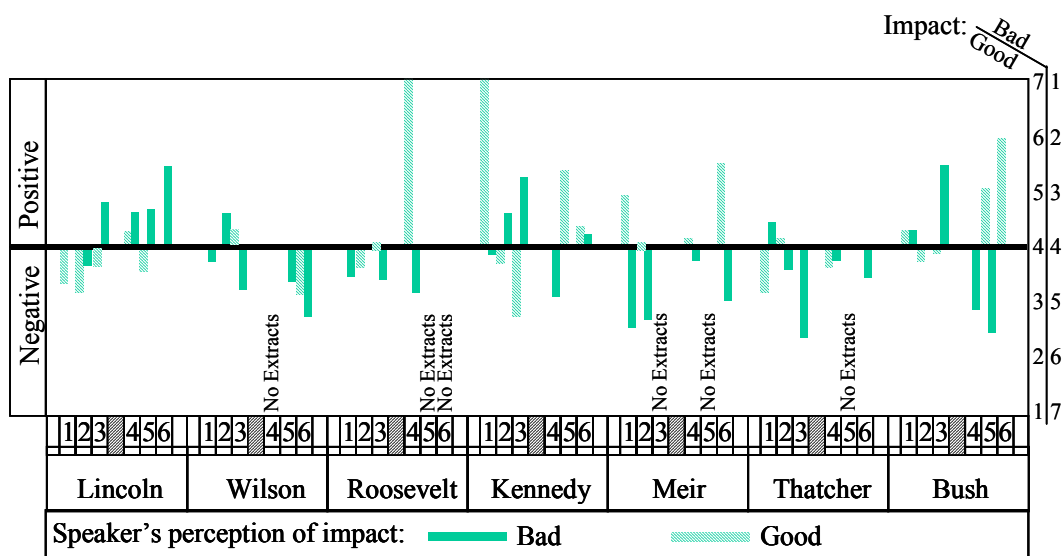


Figure 12. Explanatory style globality versus specificity component for 42 vision statements.

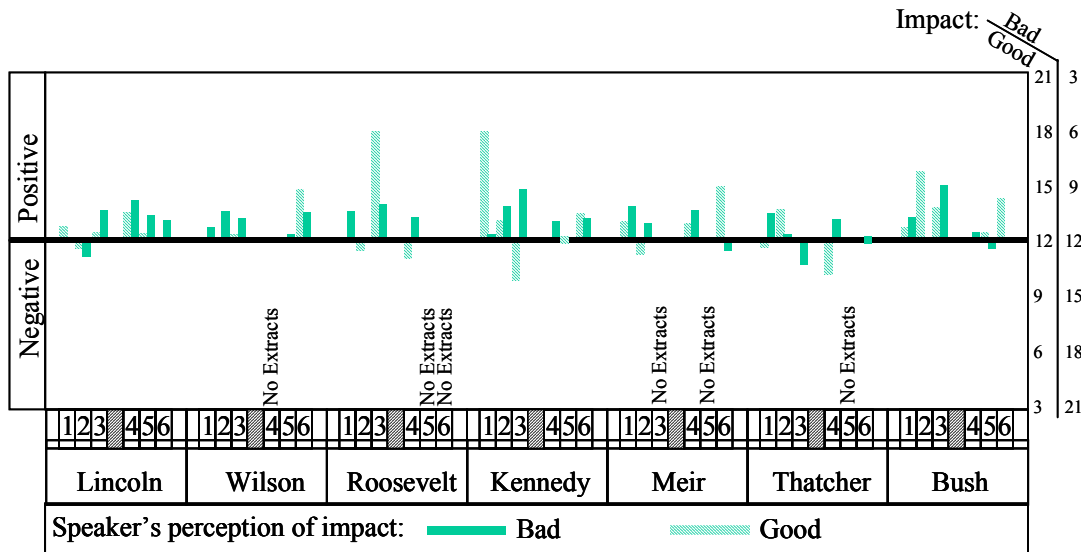


Figure 13. Explanatory style combined scores for 42 vision statements.

Reliability

I tested intercoder reliability for each of the three components of explanatory style that the raters scored. Cronbach's alpha was quite consistent: 0.5808 for internality versus externality, 0.5795 for stability versus instability, and 0.5717 for globality versus specificity. This indicates the raters were moderately consistent in interpreting how these extracts fit the scoring criteria. I conducted three significance tests of the 423 extractions using a one-sample *t*-test to determine if the means differ from a hypothetical population value. The three tests—for internality versus externality, for stability versus instability, and for globality versus specificity—all showed a *p* of < 0.0005, indicating that the null hypothesis is rejected at the .01 level.

When separated into scores for good occurrences and bad occurrences, significance tests using a one-sample *t*-test showed the results in Table 6.

Table 6 indicates that stability versus instability (bad) scores are not significantly different until they are more than 0.3 points apart, internality versus externality (bad) and globality versus specificity (bad) scores are not significantly different until they are more than 0.4 points apart, stability versus instability (good) scores are not significantly different until they are more than 0.5 points apart, globality versus specificity (good) scores are not significantly different until they are more than 0.6 points apart, internality versus externality (good) and combined (bad) scores are not significantly different until they are more than 0.7 points apart, and combined (good) scores are not significantly different until they are more than 1.2 points apart.

Table 6

One-Sample t-Test Results

Characteristic	Level	<i>p</i>	Mean	99% Confidence Interval
Internality/Externality (Bad)	.01	<.0005	2.8094	2.5890 to 3.0298
Stability/Instability (Bad)	.01	<.0005	3.6277	3.4589 to 3.7965
Globality/Specificity (Bad)	.01	<.0005	4.2641	4.0443 to 4.4839
Combined (Bad)	.01	<.0005	10.7013	10.3349 to 11.0677
Internality/Externality (Good)	.01	<.0005	4.2469	3.9204 to 4.5734
Stability/Instability (Good)	.01	<.0005	4.3958	4.1260 to 4.6656
Globality/Specificity (Good)	.01	<.0005	3.9351	3.6314 to 4.2388
Combined (Good)	.01	<.0005	12.5776	11.9942 to 13.1611

Significant Findings

Previous research created a reasonable expectation that explanatory style would be positive before a catastrophe and show more negativity afterward. Research that focused on situational change found political candidates more positive before election than afterward (Zullo, 1995; Zullo & Seligman, 1990). Other research found that candidates who were elected President of the United States from 1900 to 1984 delivered nomination acceptance speeches that scored more positive than their opponents' with the exception of Franklin Roosevelt's three re-elections. Investigators opined that Roosevelt's environment of severe economic depression and global war may have caused him to be less positive (Zullo, 1995). So it is reasonable to conclude that if explanatory style responds to situations, and if Roosevelt modeled response to bad occurrences by showing more negative explanatory style as reaction to stressful environmental influences, then the vision statements of leaders in this study should show more negative explanatory style after a catastrophe.

Raters' scoring of these 7 pivotal leaders' vision statements showed two outcomes that appear significant. In summary, the first can be seen as a norm: combined scores were predominantly neutral or in the positive direction. The second outcome is a deviation in the expected direction: 4 leaders (Roosevelt, Kennedy, Thatcher, and Bush) changed significantly toward the negative pole after a catastrophe.

The norm for the 36 scored vision statements' combined scores is that 53 (91.4%) of the 58 actual combined scores were in the positive direction or not significantly different from the neutral median, and only 5 (8.6%) were in the negative direction and significantly different from the neutral median. The implication is that the norm for the vision statements was to present an "even keel," neutral or positive, and it was a rare deviation to convey negativity. This tends to triangulate the previously reported finding that moral reasoning stage scores suggested an "even keel" appearance. The five negative visions are Lincoln on tariffs, Kennedy on Meredith, Thatcher on government

fiscal restraint, Thatcher on invasion of the Falkland Islands, and Bush declaring war on terrorism.

The deviation in the expected direction appears in collapsed data for bad occurrences for Roosevelt, Kennedy, Thatcher, and Bush, who changed significantly toward the negative pole after a catastrophe (Figure 17). A paired-comparisons *t*-test showed a *p* value of 0.020, indicating the null hypothesis is rejected at the .05 level. The other three leaders showed change toward the positive pole but not significantly; a paired-comparison *t*-test showed a *p* value of 0.140, indicating the null hypothesis is rejected at the .05 level. The shift also appeared, but not as clearly, when visions were scored as individual speeches. A paired-comparison *t*-test on the individual speech data for bad occurrences for the 4 leaders showed a *p* value of 0.049, indicating the null hypothesis is rejected at the .05 level.

Figures 14-17 graphically display the results of collapsing data into two groups that consolidate the pre-catastrophe vision statements and consolidate the post-catastrophe visions. Figure 14 shows internality versus externality, Figure 15 shows stability versus instability, Figure 16 shows globality versus specificity, and Figure 17 shows combined scores. Each figure shows two ratings for each speech if both bad and good impact extractions were available.

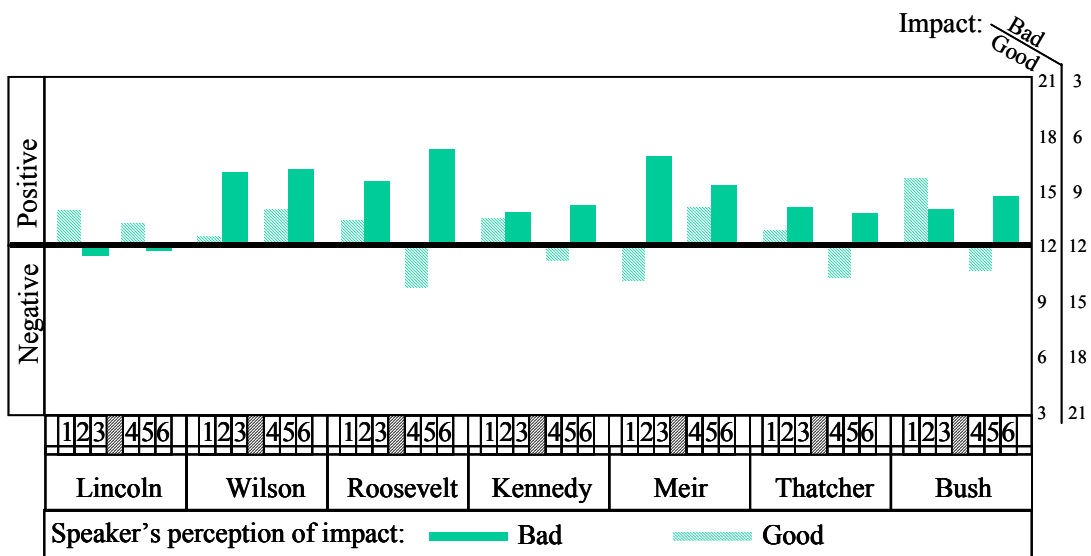


Figure 14. Explanatory style internality versus externality component for 42 vision statements aggregated by period, pre- and post-catastrophe.

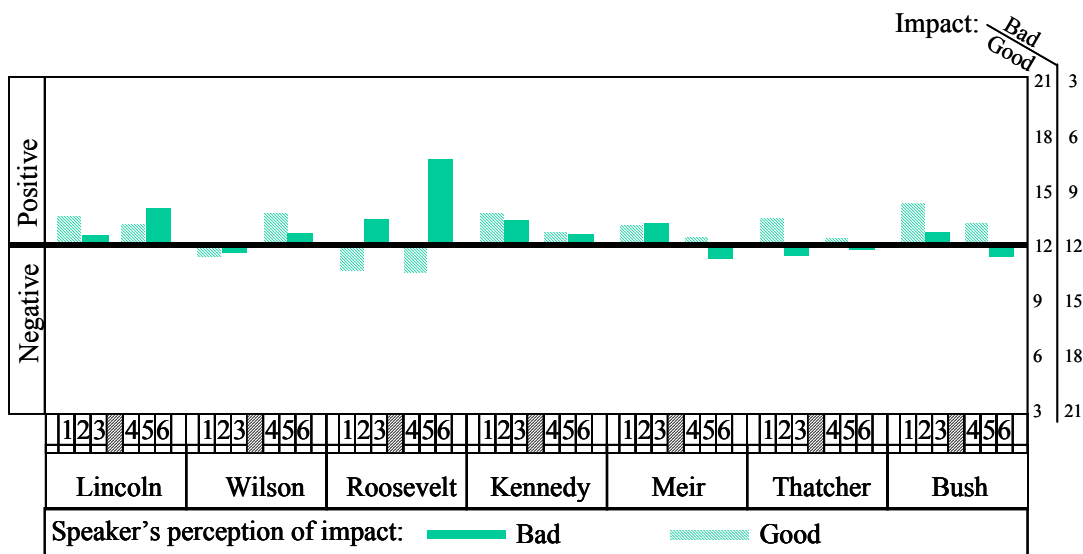


Figure 15. Explanatory style stability versus instability component for 42 vision statements aggregated by period: pre- and post-catastrophe.



Figure 16. Explanatory style globality versus specificity component for 42 vision statements aggregated by period: pre- and post-catastrophe.

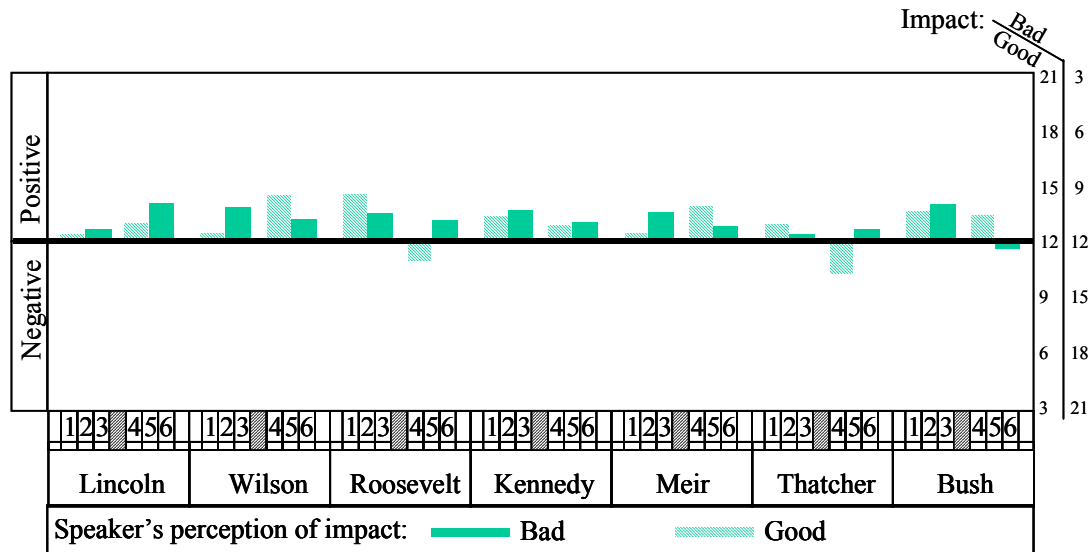


Figure 17. Explanatory style scores consolidating three components and preserving distinction between good and bad impact.

Anomalies

Six vision statements (14.3%) contained no text meeting criteria for extraction. The criteria are: a discrete occurrence, that the leader perceives to have good or bad impact, a causal explanation by the leader of what preceded the event and covaried with it, both the occurrence and explanation must be bad or both must be good, and the explanation must have a clear causal relationship to the occurrence. Four of the six visions lacked a discrete occurrence (for example, Wilson characterized the "recent course of the Imperial German Government to be in fact nothing less than war"; Wilson, 1917/1983) and two lacked a clear causal relationship (for example, both Meir and Thatcher chronicled wartime events without explaining what preceded and covaried with those events; Meir, 1973c; Thatcher, 1982b).

A question raised is whether the definition of explanatory style might privilege explanations that take the form of "us/them." The component most susceptible to an "us/them" bias is the internality versus externality component. Figure 18 shows the frequency distribution for all internality versus externality ratings. Internality versus externality scores toward 1 indicate attribution to someone or something completely external to self ("them") and scores toward 7 indicate attribution to a characteristic solely internal to self ("us," me or my country or my organization). The data show more than twice as many external scores as internal scores. What impact might this have on findings? For explanatory style (good), excess low scores on the internality versus externality component would skew scores toward the negative pole. For explanatory style (bad), excess low scores on the same component would skew

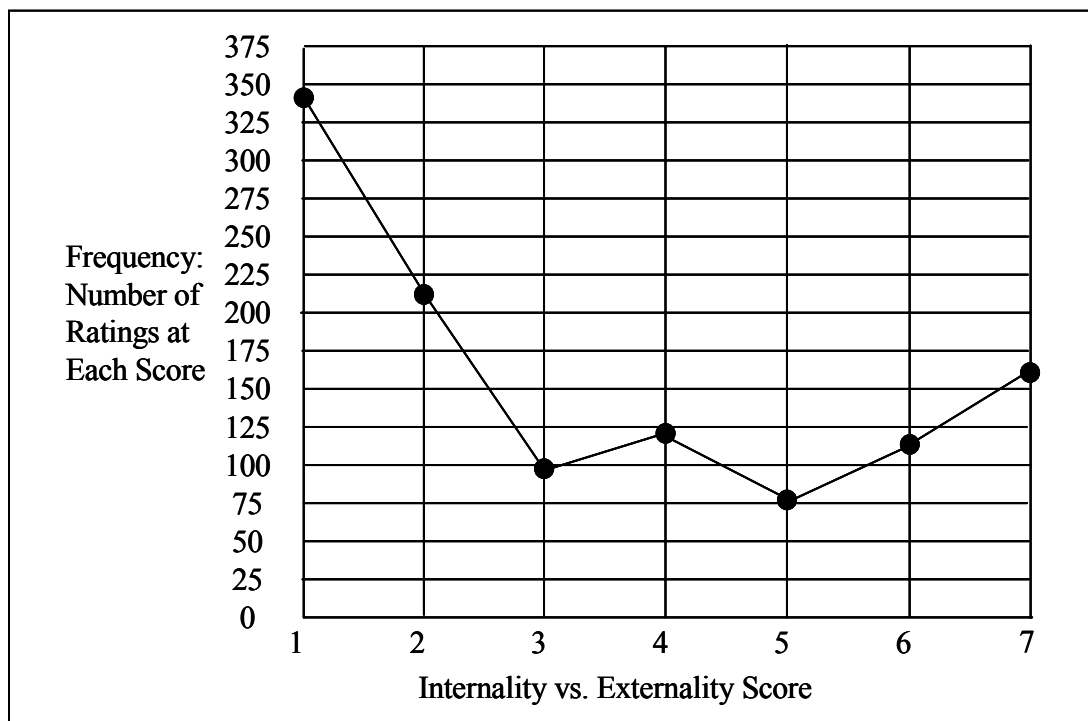


Figure 18. Frequency of ratings for internality versus externality component of explanatory style.

scores toward the positive pole. Figure 10 shows the scores for the 42 vision statements and an anomaly is that explanatory style (good) scores toward the negative pole are more visible than explanatory style (bad) scores are toward the negative pole. Also, few explanatory style (bad) scores enter the negative side of the neutral median.

Were the raters *undeservedly* biased? One triangulating test is available. Assuming that undeserved bias would cause the raters to score bad events consistently toward one of the poles, either the external ("them") or the internal ("us"), then the measure of bad events like rumination, which is a measure of the percentage of bad occurrences in the same texts, should correlate with the internal versus external component of explanatory style.³ Table 9 and Figure 22 show that rumination and internality versus externality (bad) had a negative correlation of -0.423 , significant at the .05 level with a p of 0.018, and rumination and internality versus externality (good) had no significant correlation. While this relationship bears watching, the tentative conclusion is that the ratings are valid and not significantly biased.

Another face of the internality versus externality, us versus them, issue was

³ Rumination and explanatory style for bad events usually display no correlation or low positive correlation (Zullow, 1995).

reported by Peterson (2000): "the internality dimension has become of less interest . . . [it] may well conflate self-blame and self-efficacy." He suggested that a more valid measure may emphasize only stability and globality.

Are the significant outcomes robust, or are they possibly distorted by extremes? As noted above, because some visions have more scorable extractions, a question is raised whether they might overwhelm visions with three or four extractions? For all four measures, the conservative process does not give more weight to a long vision or a vision with more extractions. Each vision is rated equally; each vision has a weight of one. For explanatory style, a test of the influence that visions with many extractions have can be conducted using a paired-comparison *t*-test to determine if the mean of casewise differences between visions with few extractions and those with many extractions differs from zero. The median number of explanatory style extractions is 10 (four have 10 extractions). Of the 36 scorable vision statements, 16 have nine extractions or less and 16 have 11 extractions or more. If the four with 10 extractions are split so the first two are added to the low-extractions set and the last two are added to the high-extractions set, then the mean of casewise differences between the low-extractions set and the high-extractions set can be tested. The paired-comparison *t*-test for that split shows a *p* of 0.052, indicating the null hypothesis is accepted at the .05 level. I conclude that the evidence shows the visions with many scorable extracts do not change the significant results for explanatory style.

Summary and Hypothesis Testing

Generally, explanatory style scores for this group of 7 leaders were neutral or positive, rarely negative. Four of the leaders showed significant changes toward the negative pole when data were collapsed into two groups, one consolidating pre-catastrophe data and the other consolidating post-catastrophe data. A shift toward negativity was expected based on previous research. The shift also appeared when visions were scored as individual speeches, but not as clearly.

Hypothesis 3 was that vision statements before a catastrophe will show some positive explanatory style and vision statements after a catastrophe will show discrepancies in the direction of more negative explanatory style.

Analysis of the vision statements used for this study partially supported the hypothesis. Overall, the 36 scored vision statements were predominantly (91.4%) neutral or positive, but 4 leaders showed significant discrepancies toward negative explanatory style in post-catastrophe data. Three other leaders showed no significant change.

Rumination

The Data Collected

As part of the 1-day scoring session, 4,400 scorable extractions from the 42 vision statements were submitted in random order to a panel of nine raters blind to each others' scoring and to possible outcomes of the study. Each extraction was scored by

three raters. The 4,400 extractions comprised all sentences in the 42 vision statements. Rumination scoring tasks raters to judge each sentence for presence or absence of rumination. For each vision statement, the sentences are recombined and the percentage of sentences containing rumination becomes the score for that vision statement.

Appendix H displays the numeric data: number of sentences in each vision statement and the rumination percentage scores. The number of sentences in a vision statement ranged from 9 to 305, with a mean of 104.7, a median of 102.5, and a standard deviation (population) of 75.18. The rumination scores ranged from 0% to 76.92%, with a mean of 39.98%, a median of 37.47%, and a standard deviation (population) of 18.26. Figure 19 displays the scores graphically.

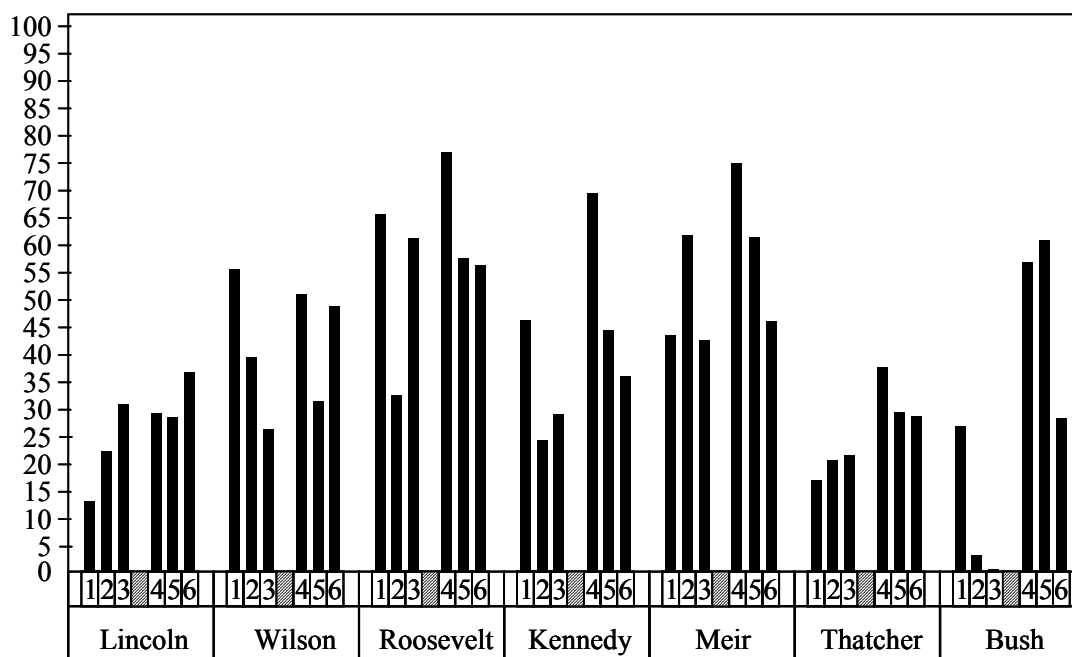


Figure 19. Percent rumination by vision, 1-3 before catastrophe and 4-6 afterward.

Reliability

Cronbach's alpha is 0.5757, the best of this study, but still indicating these raters were only moderately consistent in interpreting how these extracts fit the scoring criteria. Significance testing of the 42 scored vision statements using a one-sample *t*-test to determine if the mean differs from a hypothetical population value shows a *p* of < 0.0005, indicating that the null hypothesis is rejected at the .01 level. The mean is 39.9750.

Significant Findings

Previous research created a reasonable expectation that vision statements before a catastrophe will show less rumination than vision statements afterward. Research into rumination displayed by political candidates showed that Franklin Roosevelt showed more rumination than the opposition candidate during three re-election campaigns. The investigators opined that environmental influences may have affected Roosevelt because his presidential era, from start to finish, faced crises including severe economic depression and global war such that "ruminating about problems may have been necessary" (Zullo, 1995, p. 194).

Raters' scoring of these 7 pivotal leaders' vision statements showed four outcomes that appear significant. In summary, the first two can be seen as norms: Most of the leaders showed significant increases in rumination after a catastrophe, and for most of the leaders the vision immediately before a catastrophe contained less rumination than the vision immediately afterward. In individual speech data (Figure 19), Bush stands out for very little rumination in vision two and none in vision three. In data consolidated by period, pre- or –post catastrophe (Figure 20), Roosevelt stands out for having the most rumination and showing the least variation in amount of rumination (Table 7).

Table 7

Pre- and Post-Catastrophe Rumination Means Evaluated

Leader	Mean Sentences With Rumination	Total Sentences in Vision	Rumination Percentage
Lincoln, Pre	57	216	26.39%
Lincoln, Post	135	455	29.67%
Wilson, Pre	51	137	37.23%
Wilson, Post	119	256	46.48%
Roosevelt, Pre	170	278	61.15%
Roosevelt, Post	116	193	60.10%
Kennedy, Pre	113	335	33.73%
Kennedy, Post	140	310	45.16%
Meir, Pre	194	414	46.86%
Meir, Post	161	282	57.09%
Thatcher, Pre	146	741	19.70%
Thatcher, Post	140	439	31.89%
Bush, Pre	16	96	16.67%
Bush, Post	138	248	55.51%

Statistically, the norm is established by all 7 leaders. Their mean rumination before a catastrophe was 34.5329% and the mean after a catastrophe was 46.5571%, a

difference of 12.0 percentage points. Significance testing of the 42 vision statements using a paired-comparison *t*-test to evaluate the rumination means before and after a catastrophe showed a *p* value of 0.047, indicating the null hypothesis is rejected at the .05 level. Rumination does increase significantly after a catastrophe for this group of 7 leaders.

However, inspection of Table 7 shows only 6 (85.7%) of the leaders showed increases in rumination after a catastrophe; Roosevelt showed a decrease so small that it is essentially not a significant change in rumination level. When Roosevelt is excluded as an outlier, significance testing of the 36 vision statements of the other 6 leaders using a paired-comparison *t*-test to evaluate the rumination means before and after a catastrophe showed a *p* value of 0.039, indicating the null hypothesis is rejected at the .05 level.

Inspection of Table 7 also shows that the largest amount of change is attributed to Bush. When both Roosevelt and Bush are excluded as outliers, significance testing of the 30 vision statements of the other 5 leaders using a paired-comparison *t*-test to evaluate the rumination means before and after a catastrophe showed a *p* value of 0.004, indicating the null hypothesis is rejected at the .01 level.

Significance testing of all 7 leaders' 14 vision statements paired to compare just the vision immediately before the catastrophe with the vision immediately after it shows a *p* value of 0.010, indicating that the null hypothesis is rejected at the .01 level. This indicates that rumination does increase significantly between the vision before a catastrophe and the vision immediately afterward. The mean for vision statements immediately before a catastrophe is 30.2757% and the mean for vision statements immediately afterward is 56.7814%, a difference of 26.5 percentage points.

However, inspection of Appendix H shows only 6 (85.7%) of the leaders showed increases in rumination when the vision statement immediately following a catastrophe is compared to the vision statement immediately before the catastrophe; Lincoln showed a decrease so small that it is essentially not a significant change in rumination level. Significance testing of the 12 vision statements of the other 6 leaders using a paired-comparison *t*-test to evaluate the rumination means before and after a catastrophe showed a *p* value of 0.005, indicating that the null hypothesis is rejected at the .01 level.

A logical cause for the increase in rumination is that when something bad happens—a catastrophe—people talk about it and its effects.

In individual speech data (Figure 19), Bush stands out for very little rumination in vision two and none in vision three. A logical cause for so little rumination is that both visions advocated education reforms and looked more toward future benefits than toward past wrongs. If Schein (1992) was right about people not paying attention to a leader's vision unless disconfirming information makes them hurt, then the low amount of rumination in these two vision statements may have decreased their effectiveness. Triangulating information from before-and-after polls would be interesting to see.

In data consolidated by period, Roosevelt stands out for having the most

rumination and showing the least variation in amount of rumination (Figure 20). Table 7 shows the results of collapsing data to consolidate the pre-catastrophe vision statements into one group and the post-catastrophe visions into another group. The pre-catastrophe rumination score for Roosevelt was 61.15% and the post-catastrophe score was 60.10%,

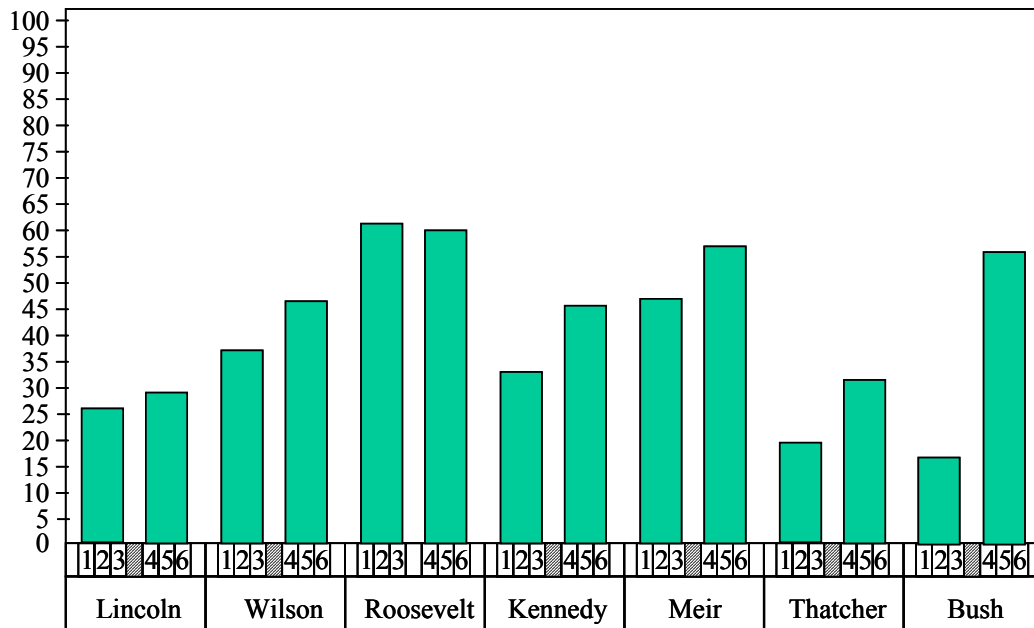


Figure 20. Percent rumination for vision statements aggregated by period: pre- or post-catastrophe.

a decrease of 1.05 points, too small to be significant. All other leaders showed increases. Significance testing of the collapsed data using a paired-comparison *t*-test showed a *p* value of .047, indicating the null hypothesis is rejected at the .05 level. A possible cause of Roosevelt's lack of change is that his presidential era faced crises including The Great Depression and World War II, logically sources of enormous long-term stress for Roosevelt and the nation and events that a president would need to discuss. The "Fireside Chats" that Roosevelt conducted by national radio broadcast stand as triangulating evidence that Roosevelt felt need to discuss the nation's problems with the nation's people. His consistently high rumination level is consistent with deliberate effort to increase "disconfirming information" in the foreground of people's thoughts and thereby win increased attention to his visions and, presumably, support for them. It is a link between empathy and pro-social visions. The disconfirming information, offered with empathic style, encourages people to support a prosocial vision that promises to remedy the situation.

Anomalies

Are the significant outcomes robust, or are they possibly distorted by extremes? As noted above, because some visions have more scorable extractions, a question is raised whether they might overwhelm visions with three or four extractions? For all four measures, the conservative process does not give more weight to a long vision or a vision with more extractions. Each vision is rated equally; each vision has a weight of one. For rumination, a test of the influence that visions with many extractions have can be conducted using a paired-comparison *t*-test to determine if the mean of casewise differences between visions with few sentences and those with many sentences differs from zero. The median number of sentences is 102.5. Of the 42 vision statements, 21 have 102 sentences or less (designated the low set) and 21 have 103 sentences or more (designated the high set). The paired-comparison *t*-test for that split shows a *p* of 0.994, indicating the null hypothesis is accepted at the .05 level. I conclude that the evidence shows the visions with many scorable extracts do not change the significant results for rumination.

Summary and Hypothesis Testing

Generally, rumination scores for this group of 7 leaders showed increases in rumination after a catastrophe. Bush stood out for very little rumination in vision two and none in vision three, and Roosevelt stood out for having the most rumination and showing the least variation in amount of rumination.

Hypothesis 4 was that vision statements before a catastrophe will show less rumination than vision statements after a catastrophe. Analysis of the vision statements used for this study partially supported the hypothesis.

Interactions

No changes after catastrophe affected all 7 leaders in the same direction. A search for patterns and trends showed three interactions to test. First, the strongest patterns in these data are the norms: explanatory style strongly neutral or positive (91.4%), no change in moral reasoning stage (71.4%), an increase toward justice in moral reasoning orientation (71.4%), an increase in rumination (85.7%). The similarities suggest examination of possible co-variation between explanatory style and moral reasoning stage. A second co-variance possibility exists between moral reasoning orientation and rumination. Third, while detailed review of the previously discovered exceptions to the norms and anomalies in the data suggested no other possible co-variations, previous research reported in the literature suggested a possible relationship between rumination and explanatory style.

Table 8 shows the results of systematically testing each of the four measures for correlation with each other using paired-comparison *t*-tests. Scatterplots are shown in Figure 21. The results show just one significant correlation. Moral reasoning orientation (caring) and moral reasoning orientation (justice) have a correlation of -1.0 with *p* <

.0005, as expected, because raters faced a polar choice: either caring or justice.

Explanatory Style and Moral Reasoning Stage

Table 8 shows the results of paired-comparison *t*-tests of the relationship between explanatory style and moral reasoning stage. The correlation for explanatory style (good) and moral reasoning stage in these data was -0.005 with a *p* value of 0.980, and the correlation for explanatory style (bad) and moral reasoning stage was -0.022 with a *p* value of 0.907. This indicates no significant correlation for both relationships.

Moral Reasoning Orientation and Rumination

Table 8 shows the results of paired-comparison *t*-tests of the relationship between moral reasoning orientation and rumination. The correlation for moral reasoning orientation (caring) and rumination was -0.289 with a *p* value of 0.071, and the correlation for moral reasoning orientation (justice) and rumination was 0.287 with a *p* value of 0.073. This indicates no significant correlation for both relationships.

Table 8

t-Test Paired-Comparison Correlation Matrix for Moral Stage, Moral Orientation, Explanatory Style, and Rumination

	Moral Stage	Moral Orientation, Caring	Moral Orientation, Justice	Explanatory Style, Good	Explanatory Style, Bad	Rumination
Moral Stage	NA	NA	NA	NA	NA	NA
Moral Orientation, Caring	Correlation -0.223 $p=0.167$	NA	NA	NA	NA	NA
Moral Orientation, Justice	Correlation 0.222 $p=0.168$	Correlation -1.000 $p<.0005^{**}$	NA	NA	NA	NA
Explanatory Style, Good	Correlation -0.005 $p=0.980$	Correlation 0.126 $p=0.547$	Correlation -0.125 $p=0.553$	NA	NA	NA
Explanatory Style, Bad	Correlation -0.022 $p=0.907$	Correlation -0.302 $p=0.105$	Correlation 0.304 $p=0.102$	Correlation 0.140 $p=0.533$	NA	NA
Rumination	Correlation -0.047 $p=0.771$	Correlation -0.289 $p=0.071$	Correlation 0.287 $p=0.073$	Correlation 0.021 $p=0.916$	Correlation 0.077 $p=0.681$	NA

****Correlation is significant at the 0.01 level.**

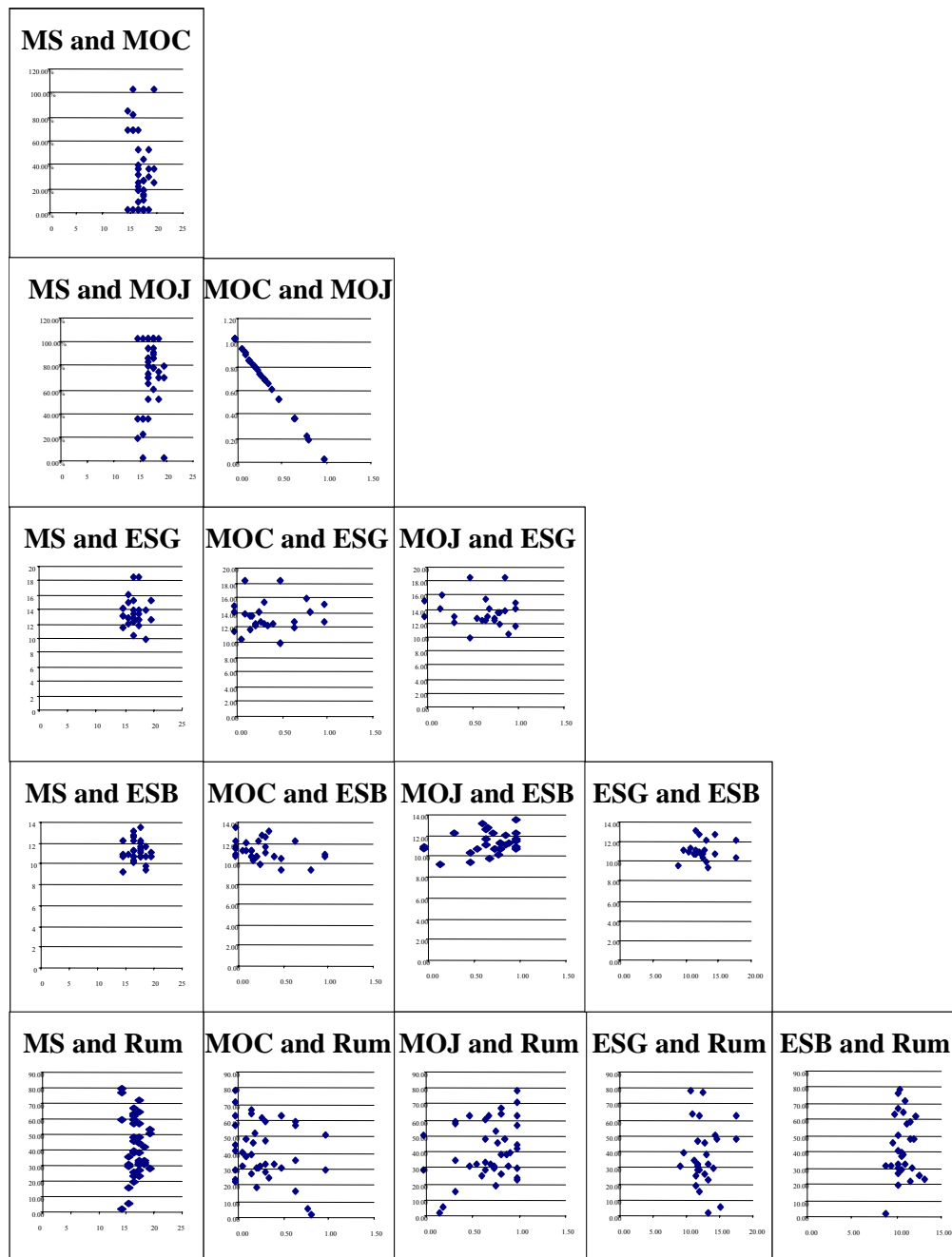


Figure 21. Scatterplots of Moral Reasoning Stage (MS), Moral Reasoning Orientation – Caring (MOC), Moral Reasoning Orientation – Justice (MOJ), Explanatory Style – Good (ESG), Explanatory Style – Bad (ESB), and Rumination (Rum).

Rumination and Explanatory Style

Table 8 shows the results of paired-comparison *t*-tests of the relationship between rumination and explanatory style. The correlation for rumination and explanatory style (good) was 0.021 with a *p* value of 0.916, and the correlation for rumination and explanatory style (bad) was 0.077 with a *p* value of 0.681. This indicates no significant correlation for both relationships.

Another set of systematic paired-comparison *t*-tests examined the relationship of rumination to each of the three components of explanatory style. Table 9 shows numeric data and Figure 22 shows scatterplots.

Table 9

t-Test Paired-Comparison Correlation Matrix for Rumination and Explanatory Style Components

	ES Combined, Good	ES Combined, Bad	ES I/E Good	ES I/E Bad	ES S/I Good	ES S/I Bad	ES G/S Good	ES G/S Bad
Rum	Correlation 0.021 <i>p</i> =0.916	Correlation 0.077 <i>p</i> =0.681	Corr. -0.279 <i>p</i> =0.159	Corr. -0.423 <i>p</i> = 0.018*	Corr. -0.126 <i>p</i> =0.532	Corr. -0.53 <i>p</i> =0.778	Corr. 0.455 <i>p</i> =0.017	Corr. 0.507 <i>p</i> = 0.004**

**Correlation is significant at the 0.01 level.

*Correlation is significant at the 0.05 level.

These tests showed two significant correlations. First, rumination and the explanatory style component internality versus externality for bad occurrences showed a correlation of -0.423 at the .05 level (*p* = 0.018). Second, rumination and the explanatory style component globality versus specificity for bad occurrences showed a correlation of 0.507 at the .01 level (*p* = 0.004). Those correlations are consistent with research in the literature that explored the combination of explanatory style bad and rumination as a variable labeled pessimistic rumination (Zullow, 1988, 1995; Zullow & Seligman, 1990). Rumination did not show significant correlations to other components of explanatory style or to the combined scores.

Patterns Observed in the Data

Moral reasoning stage, moral reasoning orientation, explanatory style, and rumination all appear to operate independently. The data also indicate that explanatory style for good occurrences operates independently of explanatory style for bad occurrences, as posited in the explanatory style scoring manual (Peterson et al., 1992). Zullow (1995) previously reported evidence of a relationship between rumination and

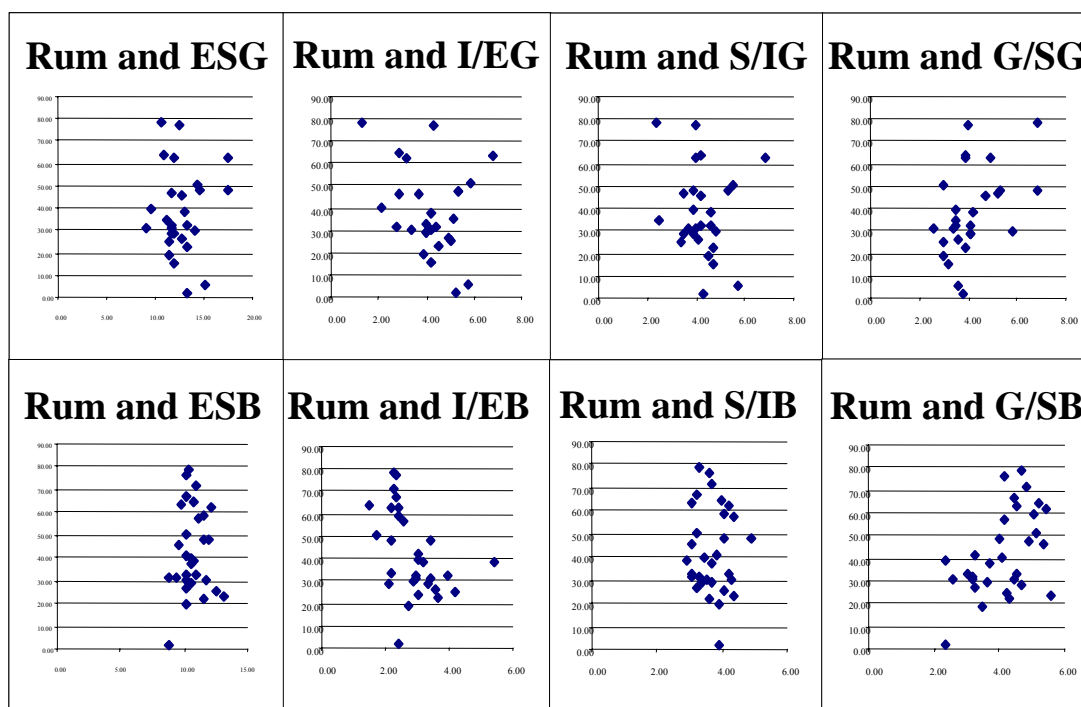


Figure 22. Scatterplots of Rumination (Rum) versus Explanatory Style – Good (ESG), Explanatory Style – Bad (ESB), and Explanatory Style components (Good and Bad) Internality versus Externality (I/EG and I/EB), Stability versus Instability (S/IG and S/IB), and Globality versus Specificity (G/SG and G/SB).

explanatory style for bad occurrences but did not indicate it was related to the internality versus externality component. In this study, rumination does show correlation to the internality versus externality (bad) (I/EB) component and to the globality versus specificity (bad) (G/SB) component of explanatory style.

Paired-comparison *t*-tests were used to compute the correlation of I/EB to G/SB. The correlation was -0.627 with $p < 0.0005$, indicating correlation significant at the .01 level. A test to compute the correlation of internality versus externality (good) (I/EG) to globality versus specificity (good) (G/SG) showed a correlation of -0.249 with $p = 0.211$, indicating no significant correlation.

One hypothesis to explain the relationship between I/EB and G/SB looks to the scoring protocol. Internality scores, by protocol, are lower when speakers attribute blame or credit to something external to self and higher when they attribute cause to self. Globality scores, by protocol, are higher when a cause affects many areas of functioning and lower when a cause affects only one area. Thus, the 7 leaders in this study may have seen causes external to self as affecting many areas of functioning (e.g., Germans causing global war) and at the same time have seen causes internal to themselves as

tending to affect only one area (e.g., I will treat you fairly).

Evidence to support the hypothesis can be found in the vision statements. An exploratory search (not exhaustive) found examples of low I/EB scores paired with high G/SB scores about war and terrorism perceived as affecting the world in vision statements by Wilson, Roosevelt, and Bush. High I/EG scores with low G/SG scores about democracy, research, and education were found in Lincoln, Wilson, and Kennedy.

Factor Analysis

I used the SPSS factor analysis procedure to search for relationships among the variables. At start, I posted raters' mean scores to a matrix that listed the 42 vision statements in chronological order as row headings and 11 variables as column headings: rumination, explanatory style good combined scores, explanatory style good internality/externality, explanatory style good stability/instability, explanatory style good globality/specificity, explanatory style bad combined scores, explanatory style bad internality/externality, explanatory style bad stability/instability, explanatory style bad globality/specificity, moral reasoning orientation, and moral reasoning stage. Because the 42 vision statements are naturally stratified by leader, when filling missing data I adhered to the stratification by using means computed from other visions in that leader's six-vision set. All analyses used principal components analysis with varimax orthogonal rotation.

For this analysis, $N = 42$ is quite small. The first analysis identified four variables that loaded on two or more factors and I excluded them from future iterations. The second analysis identified and excluded a fifth variable that loaded on two or more factors. The third iteration identified six surviving variables that loaded on only one factor (Table 10). Four factors emerged: explanatory style good (eigenvalue = 1.832, 29.796% of variance), explanatory style bad (eigenvalue = 1.622, 28.176% of variance), internality/externality for bad events (eigenvalue = 1.070, 18.215% of variance), and moral reasoning stage (eigenvalue = 1.013, 17.114% of variance). Two of the principal measures used in this study, moral reasoning orientation and rumination, loaded on all factors.

With respect to the principal measures used in this study, factor analysis showed explanatory style for good events, explanatory style for bad events, and moral reasoning stage loaded on separate factors, while rumination and moral reasoning orientation affected all of the factors. Triangulating information showed these national leaders ruminated about both good occurrences and bad occurrences, and principally used justice rather than caring moral reasoning orientation. Emergence of internality/externality for bad events as a factor was unexpected. This is the element of explanatory style that Peterson (2000) was concerned about due to possible conflation of self-blame and self-efficacy. Its emergence as a factor is consistent with the finding reported later in Table 12 that internality/externality is the one component of explanatory style that did not show correlation with the composite score.

Table 10
Factor Loadings

Components	Factors			
	Explanatory Style for Good Events	Explanatory Style for Bad Events	Internality/Externality for Bad Events	Moral Reasoning Stage
Explanatory Style Good Combined Scores	0.928			
Explanatory Style Good Stability/Instability	0.928			
Explanatory Style Bad Stability/Instability		0.919		
Explanatory Style Bad Combined Scores		0.918		
Explanatory Style Bad Internality/Externality			0.987	
Moral Reasoning Stage				0.995

Effect Size

Effect size is a power analysis quantifying the impact of a particular intervention. For this analysis, the catastrophe was used as the intervention. Post-catastrophe scores for all leaders were used to calculate the "experimental group" mean and standard deviation, and pre-catastrophe scores for all leaders were used to calculate the "control group" mean and standard deviation. (Missing data were filled by the mean of scores present.) Table 11 shows the calculated effect sizes: Catastrophe had no effect on moral reasoning stage; small effect on moral reasoning orientation, explanatory style good, explanatory style bad, and the surprise factor, explanatory style bad internality/externality; and catastrophe had a large effect on rumination (Becker, 1999; Ender, 2001). These findings are consistent with the results of paired-comparison *t*-tests reported earlier.

Table 11

Calculated Effect Sizes

Variable	Post-Catastrophe Mean	Post-Catastrophe Std. Deviation	Pre-Catastrophe Mean	Pre-Catastrophe Std. Deviation	Cohen's <i>d</i> Effect Size
Moral Reasoning Stage	17.43	1.40	17.38	1.17	0.039 None
Expl, Style Bad Internal/External	2.77	0.73	2.97	0.86	-0.174 Small
Explanatory Style Good	12.70	1.64	13.05	2.02	-0.190 Small
Explanatory Style Bad	11.13	0.71	10.78	1.04	0.193 Small
Moral Reasoning Orientation	0.44	0.97	0.13	0.96	0.321 Small
Rumination	47.16	15.54	32.79	17.94	0.856 Large

Summary

Tests showed no significant correlation among the principal measures: moral reasoning stage, moral reasoning orientation, explanatory style good, explanatory style bad, and rumination. Rumination did show correlation to two elements of explanatory style bad, internality/externality and globality/specificity. That correlation is consistent with research in the literature that explored the combination of explanatory style bad and rumination as a variable labeled pessimistic rumination (Zullow, 1988, 1995; Zullow & Seligman, 1990).

Factor analysis determined explanatory style for good events, explanatory style for bad events, moral reasoning stage, and internality/externality for bad events loaded on separate factors, while rumination and moral reasoning orientation loaded on all factors.

Effect size was consistent with previous analysis that rumination increased significantly after a catastrophe and other variables showed little consistent change.

CHAPTER FIVE

Discussion

Summary

The research question for this study asked about the impact of catastrophe on pivotal national leaders' vision statements when measured by four methods for assessing latent characteristics. The four methods examine moral reasoning stage, moral reasoning orientation, explanatory style, and rumination. Seven catastrophes and 7 pivotal national leaders were identified. For each leader, three vision statements made proximately before the catastrophe and three proximately after it were studied.

My interest in the topic arose at the time of the Los Angeles riots of 1992 when it became apparent that leaders need to deliver appropriate vision statements during times of catastrophe so everyone in their organization can do the right thing. Organizing a disaster communications network is a near-term, practical solution, but it leaves unanswered a huge question: What beliefs, perceptions, thoughts, and feelings underlie vision statements that leaders use to guide their organization through a human-caused catastrophe? Beyond manifest content, this study posited that moral reasoning stage, moral reasoning orientation, explanatory style, and rumination are four measurable aspects of vision speeches that will reflect changes in latent content.

In today's world, leaders must expect to reactively handle more human-caused catastrophes that touch large numbers of people. Today's world is characterized by strong tensions between global interdependence and expressions of diversity, increased interaction across physical and political boundaries, and by closer technological and economic ties. All three are demonstrated by the emergence of global terrorism, extensive inter-nation efforts to defend against it, and economic harm suffered by multiple nations due to a catastrophe in just one. Moreover, in today's world, the severe potential consequences of some catastrophes force responsible leaders to proactively identify organizations with high risk of experiencing a catastrophe and to find ways to prevent those catastrophes from occurring.

Three key definitions were used in this study. "Pivotal leaders" are formal leaders whose choices appear to set direction for their entire organization. In this study they are U.S. presidents Abraham Lincoln, Woodrow Wilson, Franklin Roosevelt, John Kennedy, and George W. Bush, and two prime ministers of other nations, Golda Meir of Israel, and Margaret Thatcher of Great Britain. "Human-caused catastrophe" is the second key term, defined from the perspective of the leader as a sudden crisis with such serious impact that the organization, or a significant portion of it, may not achieve fundamental goals or even survive and that appears to have been caused by people. In this study, the catastrophes are: Confederate capture of Fort Sumter, Germany's unrestricted submarine warfare, Japan's attack on Pearl Harbor, the Soviet Union's construction of offensive missile sites in Cuba, Egypt and Syria's attack on Israel, Argentina's attack on the Falkland Islands, and terrorists' attack on the United States. "Vision statement" is the

third key term, defined as a leader's prepared text meeting three criteria: (a) accessible to substantially all organization members or their designated representatives, (b) an image of the future organization, and (c) an attractive improvement over perceived alternatives.

The first of the four measures used in this study is moral reasoning stage, derived from Piaget and Kohlberg, using Commons' Hierarchical Complexity Scoring System (HCSS) and Dawson's computer scoring method. The second measure is moral reasoning orientation, assessed in light of Gilligan's discovery that moral reasoning can be based on caring as well as justice, Lyons' development of a method to comparatively score justice and caring, and Johnston's demonstration that adolescents may choose to use either rationale. The third measure is explanatory style as it has been adapted to analyze groups' and leaders' style for explaining the causes of good and bad events that affect them. The fourth measure is rumination, defined by Zullow as the percentage of sentences that describe thinking about bad events that have happened or are happening, the who, what, where, or how of a bad event, or an explanation of why a bad event occurred, or words expressing negative emotional state or display.

Two key assumptions underlying the study are that moral reasoning stage, moral reasoning orientation, explanatory style, and rumination are variables, and that available texts of the leaders' vision statements, obtained from official or authoritative sources of good reputation, are substantially accurate.

In reviewing the literature, I undertook to research five questions: (a) Does it matter that these national leaders' vision statements are influenced by advisors and consultants? (b) Does it matter what leadership theory lens is used for the research? (c) Do social psychology theories provide structural elements useful for this research? (d) Do developmental psychology theories provide structural elements useful for this research? (e) Are unobtrusive measures useful for this research?

The answer to the first question is that level of analysis is important. The influence of advisors and consultants confounds only for individual level analysis, for example attempting to predict a leader's behavior or to describe a leader's personality. If the study is focused at the organizational level, then advisors' influence is an integral part of what is being studied because it is logical for the formal leader of an organization to work with advisors and consultants.

The answer to the second question is that shared leadership fits organization level analysis. Other leadership theories focus on individual level analysis. Trait theories examine the individual's characteristics. Behavior theories examine the individual's behavior. Situation theories examine how individual and situation mediate each other. Meaning management theories examine how individual leaders manage meaning. But shared leadership theory examines the effects of multiple people working leadership together, which is the reality of the organizations being examined in this study.

In answer to the third question, three relevant streams of social psychology research were discussed. They are organizational culture, prosocial behavior, and explanatory style. Organizational culture identifies vision statements as cultural artifacts important to followers who are ready to pay attention because they are ruminating about

bad occurrences. Prosocial theory identifies vision statements as inherently prosocial because they intend to benefit people other than, or in addition to, the leader. Explanatory style uses a method of non-reactive analysis called Content Analysis of Verbatim Explanations (CAVE) to assess groups' and leaders' statements about bad or good occurrences.

Two relevant streams of developmental psychology research concerning moral reasoning stage and moral reasoning orientation were also reviewed. Stage theory posits that what people say can be measured for moral reasoning complexity on a scale ranging from least complex to most complex. Orientation theory posits that the rationale underlying people's moral reasoning can be categorized as either caring or justice.

In answer to the final question, unobtrusive measures fit catastrophe leadership research well. The alternative of high-impact deception to create catastrophe simulations is not acceptable ethically. The alternatives of asking participants to react to hypothetical situations, make self-reports, list their thoughts, and analyze retrospectively present the problem that participants know what they say will influence research outcomes. In contrast, unobtrusive study of archival records can examine unrepeatable real events across time, across cultures, and untainted by reactivity. A risk is investigator bias, which the four chosen methods control with a two-stage approach. Stage 1 is to identify segments of text that meet specific criteria and extract them. Stage 2 is to present extractions in random order to multiple raters blind to each others' ratings and to research outcomes.

Four hypotheses were identified for testing. Because previous studies found increases in moral reasoning stage after an inauguration, and decreases during stress, and after surprise attacks, Hypothesis 1 was that vision statements would show higher moral reasoning stages before a catastrophe than after. Because the influence various advisors exert in a group seems likely to vary under stress conditions, Hypothesis 2 was that vision statements will show both types of moral reasoning orientation, caring and justice, before a catastrophe and change afterward but in an unpredictable direction. Hypothesis 3 was that vision statements before a catastrophe would show some positive explanatory style and that vision statements after a catastrophe would shift toward more negativity. Hypothesis 4 was that vision statements would show less rumination before a catastrophe than after it.

The basic research design called for the investigator to study the 42 vision statements and identify appropriate extractions for each method. The criteria for moral reasoning orientation extractions are: (a) a real-life moral dilemma, (b) an asserted solution, (c) plus explanation of the dilemma or evaluation of its resolution. The criteria for explanatory style extractions are: (a) an occurrence that the leader perceives to have good or bad impact on the leader or the organization; (b) explanation by the leader of what preceded the event and covaried with it; (c) the event and the explanation must both be good or both be bad; (d) and a clear causal relationship must exist. The criterion for rumination is simply that every sentence in the text is extracted for randomized presentation to the raters so they see it devoid of all context. In this study, the vision

statements were prepared for computer-assisted expert scoring of moral reasoning stage by T. L. Dawson by creating a block file for each speech, devoid of paragraphing, contractions and abbreviations, and spelling out all numbers. Had these vision statements been scored manually, the criteria for extractions are: (a) an asserted solution, (b) to a problem, (c) plus justification of the solution.

Reliability of scores was addressed in two ways. For moral reasoning stage, T. L. Dawson served as an expert rater supported by tested computer software having a Cronbach's alpha equivalent equal to 80% to 97%, which is widely considered as effective reliability. For the other three measures scored manually by blind raters, I computed Cronbach's alpha values ranging from 0.4092 to 0.5757, which is only moderate reliability.

The findings for each of the four measures are summarized below.

Moral Reasoning Stage

Of the 42 vision statements, 41 were scorable by an expert rater and one had no content meeting extraction criteria. The scores were meaningfully different. A *t*-test rejected the null hypothesis at the .01 level ($p < 0.0005$). Generally, moral reasoning stage scores for this group of 7 leaders appeared to approximate a normal curve. Three significant findings were a norm of no significant change immediately after a catastrophe, and two deviations from that norm, leaders who showed a decrease or increase. Evaluated as individual speeches, the data showed a significant drop in moral reasoning stage immediately after a catastrophe for Roosevelt ($\Delta = 2$) and Meir ($\Delta = 3$) but no significant change for the other 5 leaders (71.4%). Additional data showed three leaders displayed a low post-catastrophe vision statement that was followed by another that had significantly higher moral reasoning stage ($\Delta = 2$). That possibly indicates deliberate simplification of the first speech followed by elaboration in the second. Collapsed data showed no significant drop in moral reasoning stage after a catastrophe for any of the leaders ($p = 0.879$ for all 7 leaders and $p = 0.057$ for Meir alone). Hypothesis 1 was not supported: The norm was no significant differences in moral reasoning stage before and after a catastrophe.

Moral Reasoning Orientation

Of the 42 vision statements, 40 were scorable by a nine-rater panel and two had no content meeting extraction criteria. The number of extractions scored was 115. The scores were meaningfully different. A *t*-test rejected the null hypothesis at the .01 level ($p < 0.0005$). Generally, the leaders (71.4%) showed an increase in justice orientation following a catastrophe but one leader showed an opposite change, toward caring. Evaluated as individual speeches, the weighted data showed a significant increase at the .05 level ($p = 0.044$) in justice orientation for the 5 leaders with scorable visions immediately before and after a catastrophe. Thatcher showed an opposite increase toward caring orientation. Bush dramatically changed from caring to justice orientation and the change occurred when presidential advisor Hughes withdrew from drafting his

speeches. Generally, the norm for these American presidents was to show considerable caring. The 2 non-American leaders, Meir and Thatcher, showed higher justice orientation. Why? A number of possible reasons can be imagined, for example, because they speak in a parliamentary forum, or because they are the only two women, or because circumstances forced them to display special toughness. Available data are insufficient to determine the true reason and triangulating information that would illuminate the cause remains to be located. Hypothesis 2 was partially supported: Change in moral reasoning orientation usually occurred when measured by the vision statement immediately before a catastrophe compared to the vision statement immediately after. The change usually was toward justice but once toward caring.

Explanatory Style

Of the 42 vision statements, 36 were scorable by a nine-rater panel and six had no content meeting extraction criteria. The number of extractions scored was 423. The scores were meaningfully different. A one-sample *t*-test rejected the null hypothesis at the .01 level ($p < 0.0005$). Overall, the 36 scored vision statements were predominantly (91.4%) neutral or positive, but four leaders showed significant changes in the direction of the negative pole. Hypothesis 3 was partially supported.

Rumination

All 42 vision statements were scorable and the number of sentences scored was exactly 4,400. The scores were meaningfully different. A one-sample *t*-test rejected the null hypothesis at the .01 level ($p < 0.0005$). Four outcomes appear significant: Most of the leaders showed a significant increase in rumination after a catastrophe; for most leaders the vision immediately before a catastrophe contained less rumination than the vision immediately afterward; Bush stood out for very little rumination in vision two and none in vision three; Roosevelt stood out for having the most rumination and showing the least variation. Evaluated as individual speeches, paired-comparison *t*-tests showed significant increases in rumination at the .01 level ($p < 0.0005$) after a catastrophe across all leaders and all visions. Paired-comparison *t*-tests for the data of all 7 leaders showed an increase in rumination significant at the .01 level ($p = 0.010$) when the vision immediately after a catastrophe was compared to the vision immediately before it. Collapsed data showed increases in rumination after a catastrophe for all leaders except Roosevelt, who already was higher in rumination than any other leader. Paired-comparison *t*-tests showed that the changes for collapsed data were significant at the .05 level. Hypothesis 4 was partially supported: Rumination scores for this group of 7 leaders showed an increase in rumination after a catastrophe but the increase was significant only sometimes.

Interactions

The four measures operated independently. Rumination showed correlation with two components of explanatory style for bad occurrences and those two components

were negatively correlated with each other. This may be explained by the 7 leaders perceiving that causes external to themselves affect many areas of functioning while causes internal to themselves tend to affect only one area. Evidence tending to support that hypothesis can be found in the content of vision statements. Factor analysis showed that three of the principal measures used in this study, moral reasoning stage and explanatory style for both bad and good occurrences, were underlying factors. A fourth factor also emerged, the component of explanatory style called internality/externality for bad occurrences. Two of the principal measures, moral reasoning orientation and rumination, were associated with all four of the underlying factors.

Conclusions

In drawing conclusions, I focused on methods, content, and alternative explanations.

Methods

These four methods have been broadly categorized as content analysis techniques. Within content analysis, a number of method subcategories have been recognized that are not always mutually exclusive, such as thematic, semantic, and network analysis. "Thematic," a method often encountered, is analysis of text for manifest themes or concepts; inferences focus on the words used (Smith 1992a). Semantic approaches analyze sentence segments that interrelate themes and network approaches analyze locations of interrelated themes (Roberts, 1997). "Structure" examines the elements that constitute key ideas and how those elements are assembled. Moral reasoning stage is a complex example of the structure approach because it examines element sub-assemblies to evaluate what hierarchical stage they achieve. Moral reasoning orientation and explanatory style are simpler examples of the structure approach because they assess bi-polar elements to assess which ones compose the preponderance of the text. Rumination is the simplest example because it assesses the percentage of sentences composing a text that contain rumination. The structure approach does not infer meanings from the manifest words used or the interrelationships or locations of manifest themes. "Yes, we have no bananas" could be thematically analyzed around such frames as agreement (yes), person (we), or fruit (bananas). Structure analysis might score it a sentence containing rumination (no bananas).

Two observations about intercoder reliability in this study are that Cronbach's alpha decreased from the training session to the research session (Table 12), and that intercoder reliability was only moderate at best.

Training on the rating method in advance of the research session is the primary control on rater variability. Table 12 shows quite acceptable levels of intercoder agreement during training. The decrease during the research session might be attributed to a number of causes. One is that the training texts, taken from literature prepared by experts, may be simpler or more clearly display characteristics that match scoring criteria than do the research texts that were extracted from real-world vision statements.

Table 12

Intercoder Reliability Measurements

Topic	Training Session α	Research Session α
Moral Reasoning Orientation	0.8708	0.4092
Explanatory Style, Internality Versus Externality	0.7761	0.5808
Explanatory Style, Stability Versus Instability	0.6927	0.5795
Explanatory Style, Globality Versus Specificity	0.8475	0.5717
Rumination	0.9020	0.5757

Another cause might be rater distraction due to health concerns. During the research session, two raters scored only two-thirds of the scheduled texts. One said dyslexia made the task difficult and the other reported onset of a migraine headache. Both volunteered to complete the unfinished work later at home; the rater with the migraine did so, but the rater with dyslexia never did. A third cause might be raters' choices of decision-making processes. Researchers expect raters to rationally follow the prescribed scoring rules. However, raters may substitute their own rules and perhaps are likely to do so as experience increases and they feel more educated and expert. For example, one rater who scored moral reasoning stage extracts not used in this study explained that she looked for patterns matching examples in the scoring rules, but learned to also consider the length of extracts and "commas, complex words, and transitional phrases. The ones with a higher degree of these components got a higher score" (J. Magro, personal communication, September 16, 2003). Shapiro (1997) described efforts to control raters' judgments as frustrated by lack of success for the very reason that raters are not given detailed rules to be tediously applied but instead are expected to interpret general rules and make human judgments.

The first cause might be better controlled with a larger pool of expertly scored training items drawn from comparable vision statements. The second and third causes might be better controlled by replacing the ad hoc, one-time scoring team that I used with a more permanent team composed of people who competed to fill vacant positions and whose training scores matched or increased the team's overall Cronbach's alpha on training items.

A critic might question whether it is truly desirable to maximize the scoring team's Cronbach's alpha. High scores might indicate the members are able to predict other team members' ratings or the "politically correct" rating. The ad hoc rating team that I used might be better, especially when encouraged to score independently, blind to how other raters actually score the same item, and without regard to how anyone else might score it. Such raters arguably behave more like real organization members and

how they interpret the leader's vision statement.

Overall, the variability in rater scoring supports the Webb et al. (1966) opinion that unobtrusive, non-reactive measures of real-world events do not allow the same control over variables available under laboratory conditions. Webb et al. encouraged accepting data as useful even when statistical reliability fell below levels expected in laboratory experiments. Apply judgment and intuition, and look to replication and triangulation to confirm results.

These methods for exploring latent features of vision statements worked well and should be used again. Similarities between these scores and scores by different raters on a subset of these texts during a pilot study indicate the results are robust even when intercoder reliability was only moderate or when criteria prevent finding qualified extractions in some vision statements. Variation appears to be primarily where it should be—in the leaders' vision statements.

Computer scoring, as modeled by expert scoring of moral reasoning stage, eliminated two sources of variance: tedious human identification of text meeting criteria to be extracted, and tedious human scoring of the extractions. A risk is that the human judgments are not really eliminated, just moved earlier in the process when the scoring rules are written into the program. Computer scoring then applies those rules in blanket fashion without sensible adjustments that human raters might make (Shapiro, 1997). If wise design and testing can prevent such computer mis-scoring, then availability of high quality computer scoring would allow the analysis of large numbers of texts, although that would not change this study that focused on a few texts most proximate to a catastrophe. Assuming validation has been adequate, the HCSS approach to scoring moral reasoning stage, with its ability to unobtrusively analyze any text without issues of domain for questions or issues of moral reasoning orientation for answers, stands as a model for other unobtrusive methods to emulate.

Content

This research sprang from my curiosity about unconscious beliefs, perceptions, thoughts, and feelings that underlie pivotal leaders' vision statements used to guide organizations after a human-caused catastrophe. The techniques I used are just four among many that might make latent variables visible. All four qualities vary across the six vision statements associated with each leader, rarely are they the same for two consecutive vision statements, and they vary independently.

From a systems viewpoint, these four variables may describe four levers used in leaders' vision statements to influence organization members. Organizational systems theory recognizes levers as a key principle. When attempting to change what an organization does, even massive pressure—brute force—may not work. "[K]nowing where to intervene so that a small effort can get a huge result" is identifying a lever and using it (O'Connor & McDermott, 1997, p. 21.) Organizational systems can contain many levers. What do the data disclose about the operation of these four variables as levers in these leaders' visions statements? That information would answer the research

question: What is the impact of human-caused catastrophe on pivotal national leaders' vision statements?

Moral reasoning stage. To test Hypothesis 1, I looked for a decrease in moral reasoning stage following a catastrophe. The data failed to show any consistent decrease but it did show variability for each leader. The overall frequency distribution pattern (Figure 1) approximated a normal curve, indicating fluctuation around a central tendency of about Stage 17, which is defined as using systems to organize efforts while relying on linear logic. Using systems to organize efforts well describes a role of national government. Relying on linear logic may describe a process for making decisions rationally, which seems an image that democracies would aspire to project.

For some leaders, discernible patterns did emerge. Wilson's vision statements uniquely tended to score at Stage 20, systems that could become metasystems—systems of systems. Triangulating information describes Wilson as focused on multi-national relationships and organizations that led to the League of Nations. Bush's visions tended to score at Stage 15 or 16, very linear in their logic. Triangulating information describes Bush as insistent that his speechwriters use "strict linear logic" (Frum, 2003, p. 48). Like Bush did after September 11, 2001, Roosevelt also delivered a Stage 15 vision following the Pearl Harbor attack. Triangulating information describes this as a deliberate choice to be brief and uncomplicated. Roosevelt rejected Secretary of State Cordell Hull's recommendation that the speech be a thorough recitation of Japanese treachery (National Archives and Records Administration, 2001, 2003). While the available triangulating information is not comprehensive, it suggests that for these pivotal leaders' vision statements moral reasoning stage did not vary in unreasoned reaction to events but instead was deliberately adjusted. While this might be "impression management," viewed from an organizational systems perspective it could be adjustment of a lever to influence organization members to act in the desired way. Wilson taught support for multi-national relationships instead of isolationism. Bush and Roosevelt taught support for controlled response to danger. Frum (2003) characterized Bush's contributions to the nation after the 9/11 catastrophe as calm, self-restraint, moderation, quiet determination, and lack of vengefulness or anger.

Moral reasoning orientation. With respect to Hypothesis 2, the data did show both caring and justice orientations in vision statements and a change with unpredictable direction following a catastrophe. Justice orientation clearly dominated the vision statements as a group, and the predominant change after a catastrophe was in the direction of more justice orientation. Nothing in the moral reasoning orientation data suggests variation is due to unreasoned reaction to events. Instead, the data suggest deliberate choices that could be adjustment of an organizational system lever to influence organization members. Available triangulating information supports this view. Bush's first vision after the terrorist attacks moved somewhat toward justice but remained predominantly caring. The lead speechwriter responsible for that vision statement was criticized by other staff members, a new lead speechwriter assumed responsibility, and the second vision after the attacks moved strongly toward justice.

Frum (2003) said this was because Bush changed: a peacetime commitment to compassion (caring) became a wartime commitment to fight to victory (justice). Margaret Thatcher showed little change. She supported national honor, international law, and democracy and self-determination for the Falkland Islanders without factoring in the human cost of war (Campbell, 2003). Hers were justice arguments.

An important observation in the data is that raters did see both caring and justice orientations in these vision statements, sometimes in the same extraction. Carol Gilligan originally saw caring as an alternative to the justice orientation. She described them as "two different modes . . . While an ethic of justice proceeds from the premise of equality—that everyone should be treated the same—an ethic of care rests on the premise of nonviolence—that no one should be hurt" (Gilligan, 1993, pp. 173-174). Nona Lyons (1988) and Kay Johnston (1988) both found that individuals often used both modes simultaneously but one mode predominated. Gilligan and Attanucci (1988) charted a five-step scale: care only, care focus, care-justice, justice focus, justice only. All of these analyses used a linear scale with care at one pole and justice at the other. Lyons (1988) explicitly used polar scoring: counting the number of caring items and the number of justice items and naming the mode with the highest number the predominant mode.

A critic might challenge the linear scale assumption. Framing moral orientation research on a linear model with justice at one pole and caring at the other produces results that are clear where "only caring" or "only justice" were used. But the evidence shows people use both, perhaps simultaneously. Caring and justice may not be mutually exclusive. Where raters find both appearing simultaneously, the scoring results are muddled. Framing moral orientation research on a two-axis grid may support better analysis and interpretation. One axis would show how much justice orientation raters perceived in a text, and the other axis would show how much caring orientation raters perceived in the same text. Like chocolate syrup and milk always blend but the resulting chocolate milk has characteristics that reflect the ingredients' relative proportions, so too caring and justice may blend to produce moral reasoning orientation reflecting their relative proportions. Data from this study are so plotted in Figure 23, but the charts are illustrative rather than definitive because raters were asked to score on a linear scale rather than on two separate scales. For these charts, I adopted a grid measuring 2 x 2 because the largest scalar point in the data was 1.8.

The two-axis grid is reminiscent of others associated with leadership theory, such as the Blake and Mouton (2001) managerial grid, the Hersey and Blanchard (1982) situational leadership quadrants, and the Ohio State leadership studies quadrants (Hersey & Blanchard, 1969). The Blake and Mouton axes are concern for people (similar to care) and concern for production (similar to justice in its focus on process and rules). The Hersey and Blanchard axes are supportive behavior (similar to care) and directive behavior (similar to justice in its focus on process and rules). The Ohio State axes are consideration (similar to care) and structure (similar to justice in focus on process and rules).

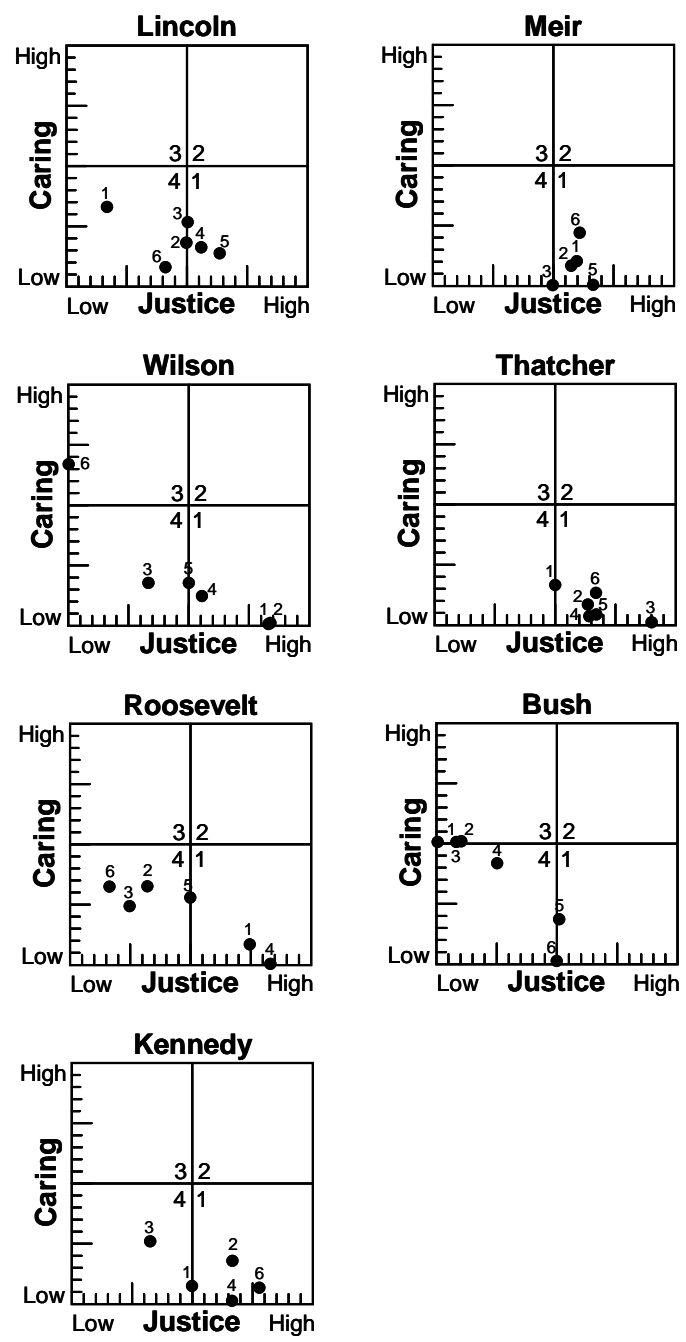


Figure 23. Moral reasoning orientation plotted on two axes (illustrative).

Grid analysis for moral reasoning orientation is an unanticipated direction for this research and not well supported by the data collected because raters were asked to score on a single linear scale instead of on two separate scales. Therefore interpretations of this moral reasoning orientation data using the similar Blake and Mouton, Hersey and Blanchard, or Ohio State grids are not attempted here.

Explanatory style. To test Hypothesis 3, I looked for more explanatory style in the negative direction after a catastrophe. The data showed some change in the negative direction, but the predominant finding was that 91.4% of the vision statements were neutral or positive. Theory underlying explanatory style recognizes this pattern as characteristic of normal people and, by extension, characteristic of normal leaders (or leader groups). The data failed to show any consistent, post-catastrophe change toward negativity that underlying theory would identify as a concern if it approached the extreme end of the negative scale.

As noted earlier, Peterson (2000) said the internality dimension may conflate self-blame and self-efficacy. I explored what impact such conflation might have on this study. Two observations are helpful. First, the validity and reliability of the CAVE technique have been demonstrated (Peterson, 1992; Peterson et al., 1995; Reivich, 1995; Schulman et al., 1989). Second, self-blame ought to appear only for bad occurrences attributed to myself, or my organization, or my nation. The 423 extractions used for this study contained only eight (0.0189%) bad occurrence extractions that raters' mean scores placed at 6.5 or more on a 7-point scale as a strong attribution to self. Examination of those eight showed only one possibly interpretable as self-blame: Lincoln said trading in foreign slaves was imperfectly suppressed and fugitive slaves were only sometimes surrendered to authorities because our citizens' moral sense imperfectly supports those laws. Two of the eight could be interpreted as self-efficacy: Lincoln chose to leave three Supreme Court judgeships vacant so he could appoint southern judges for balance when peace was restored; and Lincoln declined to use military courts to establish justice in insurgent areas because he chose to avoid using unnecessary power. The other five were neither blame nor efficacy. For example, Falkland Islanders were in tears because they did not want to be Argentine. I conclude that for this study the effect of conflation is insignificant.

Buchanan and Seligman (1995) called for analysis of the three components of explanatory style separately and comparison to the composite score. Table 13 shows those comparisons.

For bad occurrences, globality and internality correlated negatively. The vision statements attributed narrow, specific causes to self (internal) and broad, global causes to something external. For bad occurrences, globality and stability also correlated positively. The visions saw narrow, specific causes as one-time events (unstable) and broad, global causes as persistent and perhaps endless.

For good occurrences, internality and stability correlated positively. Internal causes will persist (stable) and external causes will never occur again (unstable). This pattern is consistent with previous research that found explanatory style scores for good

Table 13

Correlation of Explanatory Style Components to Each Other and to Composite Scores

	I/E and S/I	I/E and G/S	S/I and G/S
Bad Occurrences	Correlation: -.044 Significance: 0.780	Correlation: -.592** Significance: <0.0005	Correlation: .358* Significance: 0.020
Good Occurrences	Correlation: .622** Significance: <0.0005	Correlation: -.216 Significance: 0.170	Correlation: -.066 Significance: 0.678
	I/E Bad and I/E Good	S/I Bad and S/I Good	G/S Bad and G/S Good
	Correlation: .029 Significance: 0.857	Correlation: -.183 Significance: 0.247	Correlation: .261 Significance: 0.095
	I/E and Composite	S/I and Composite	G/S and Composite
Bad Occurrences	Correlation: -.092 Significance: 0.562	Correlation: .449** Significance: 0.003	Correlation: .541** Significance: <0.0005
Good Occurrences	Correlation: .186 Significance: 0.237	Correlation: .407** Significance: 0.007	Correlation: .387* Significance: 0.011

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

and bad occurrences are independent. The pattern is confirmed by absence of correlation between each component's scores for good and bad occurrences.

The stability and globality components show correlation with the composite score, indicating they are drivers shaping the composite. This is consistent with Peterson's observation that stability and globality may someday prove to be a valid measure of explanatory style in themselves.

Overall, explanatory style and its component elements showed no pattern of reaction to catastrophe but the composite score did show consistent neutral or positive scores. These would be characteristic of normal people and, by extension, characteristic of normal organizations. From a systems view, leaders' vision statements could be using this lever to signal organization members to keep an "even keel" and react without negativity. Providing triangulating information, Frum (2003) commented on Bush's second vision statement after the terrorist attack, calling attention to the most important thing that Bush's vision did *not* say: it did not accept or deign to acknowledge any argument that the United States somehow caused the terror attacks. That would have been self-blame for a bad occurrence that would score toward negativity. Similarly,

Abraham Lincoln clearly blamed bad occurrences on external causes: the Civil War was not caused by Union actions but instead by Confederate aggressors who attacked non-threatening Fort Sumter.

Rumination. For Hypothesis 4, I expected vision statements to show more rumination after a catastrophe and they did. Roosevelt, whose term in office was marked by a series of serious crises, showed high rumination both before and after the Pearl Harbor catastrophe. After a catastrophe, rumination comprised 30% to 60% of the sentences in these vision statements. If Schein (1992) was right, people are only ready to pay attention to a leader's vision when disconfirming information makes them hurt. Rumination is disconfirming information. This supports the systems view that leaders could use rumination as a lever to influence their organization members to support a vision the leader feels is important, as the vision following a human-caused catastrophe would be. Again Frum (2003) offered triangulating evidence. He criticized Bush's first vision statement after the terrorist attack for addressing a circumstance full of rumination but failing to provide a vision organization members could support. He praised Bush's second vision statement after the attack for adding to the rumination such a vision: Every nation has to decide if they are with us or with the terrorists. The United States will regard any nation harboring or supporting terrorism as a hostile regime.

Alternative Explanations

One explanation of the data in this study could be that the vision statements showed no significant change in moral reasoning stage or explanatory style following a catastrophe and that moral reasoning orientation and rumination did show significant change. This is supported by the analysis that showed these variables were independent and that justice orientation and rumination usually did increase after a catastrophe. If this were a univariate study, any one of these results might be reported as the finding. But this is a multivariate study and that interpretation does not explain why the data show, across each leader's six vision statements, variations in moral reasoning stage, moral reasoning orientation, explanatory style, and rumination.

We need something more, and an alternative explanation better fits all of the data observed. Organizational systems theory recognizes levers as a key principle that allows a small effort by the leader to achieve a large result. Systems can contain many levers. The data appear to show these leaders' vision statements adjusting these four variables to influence organization members to behave in a desired direction. This explanation is supported by the correlation and factor analysis tests that showed independence in these four variables, along with triangulating evidence that showed at least some of the changes were purposeful. This explanation answers questions about why organization members attend to leaders' vision statements, what benefit they derive, and how a leader achieves community feeling and response in the organization.

The organizational lever explanation is consistent with previous research that used explanatory style to measure optimism, pessimism, and rumination in political candidates' campaign speeches. Except for Franklin Roosevelt, the speeches of

presidential candidates who won election showed more optimism and less rumination than their opponents' speeches. Zullo also reported anecdotal evidence indicating the nomination acceptance speech made by Michael Dukakis in 1988 intentionally used optimism [as a lever] (Zullo, 1988, 1995; Zullo et al., 1988; Zullo & Seligman, 1990).

From an organizational culture viewpoint, in Schein's (1992) terms, adjusting these four characteristics through a vision statement that delivers them to organization members constitutes a fundamental behavior, perhaps even an unconscious behavior, that in the aggregate can be identified as values and that when written or spoken becomes an artifact (Schein's three-level model of culture is artifacts, traced to values, arising from foundational beliefs, thoughts, and feelings—and by implication behaviors reflecting those beliefs, thoughts, and feelings).

From a prosocial behavior viewpoint, assuming that a vision statement adjusts these four characteristics for maximum effect, to offer what organization members will see as an attractive improvement over other alternatives, then these four variables are being used to create or enhance the prosocial quality of the vision statement. By definition, "prosocial" is action intended to benefit people other than, or in addition to, one's self.

The role of empathy remains unclear. Do the levers help vision statements convey leaders' empathy to organization members? Or is it the other way around: Do the levers help the leaders' vision statements evoke empathy from organization members? Or is there a middle ground, perhaps, a climate generating or supporting empathy felt by both leaders and organization members? A view reported in the literature, that leaders use empathy to help followers adjust to environmental changes and that empathy generally is unconscious (Bass, 1990; Turner et al., 2002), fits the systems view of these levers in operation but does not clarify the direction empathy flows.

Overall, this study shows an impact of human-caused catastrophe was adjustment of these four variables in pivotal national leaders' vision statements as if they are organizational levers.

What impossible matter will he make easy next?
(Shakespeare, 1611/1997, p. 1669)

References

- Amato, P. R. (1983). Helping behavior in urban and rural environments: Field studies based on a taxonomic organization of helping episodes. *Journal of Personality and Social Psychology*, 45(3), 571-586.
- Avolio, B. J., Jung, D. I., Murry, W., & Sivasubramaniam, N. (1996). Building highly developed teams: Focusing on shared leadership processes, efficacy, trust, and performance. In M. M. Beyerlein, D. A. Johnson, S. T. Beyerlein (Eds.), *Advances in interdisciplinary studies of work teams: Team leadership* (Vol. 3, pp. 173-209). Greenwich, CT: JAI Press.
- Axelrod, R. (1976). *Structure of decision: The cognitive maps of political elites*. Princeton, NJ: Princeton University Press.
- Ballard, E. J. (1983). Canadian prime ministers: Complexity in political crises. *Canadian Psychology*, 24(2), 125-129.
- Basler, R. P. (1953). *The collected works of Abraham Lincoln* (Vols. 4-5). New Brunswick, NJ: Rutgers University Press.
- Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership: Theory, research and managerial applications* (3rd ed.). New York: Free Press.
- Bass, B. M. (1998). *Transformational leadership: Industrial, military, and educational impact*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 282-316). New York: McGraw-Hill.
- Batson, C. D., Ahmad, N., Lishner, D. A., & Tsang, J-A. (2002). Empathy and altruism. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 485-498). New York: Oxford University Press.
- Becker, L. A. (1999). *Effect size*. Retrieved December 30, 2003, from University of Colorado at Colorado Springs web site: <http://www.uccs.edu/~lbecker/psy590/es.htm>
- Bennis, W. (1999). The end of leadership: Exemplary leadership is impossible without full inclusion, initiatives, and cooperation of followers. *Organizational Dynamics*, 28(1), 71-80.
- Bennis, W., & Nanus, B. (1997). *Leaders: Strategies for taking charge* (2nd ed.). New York: HarperBusiness.
- Berkowitz, L. (1953). Sharing leadership in small, decision-making groups. *Journal of Abnormal and Social Psychology*, 48(2), 231-238.
- Binder, A. (1964). Statistical theory. In P. R. Farnsworth, O. McNemar, & Q. McNemar (Eds.), *Annual Review of Psychology* (Vol. 15, pp. 277-310). Palo Alto, CA: Annual Reviews.

- Blake, R. R., & Mouton, J. S. (2001). Grid organization development. In W. E. Natemeyer & J. T. McMahon (Eds.), *Classics of organizational behavior* (pp. 436-445). Prospect Heights, IL: Waveland Press. (Reprinted from *Personnel Administration*, January – February 1967).
- Bowers, D. G., & Seashore, S. E. (1966). Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative Science Quarterly*, 11(11), 238-263.
- Bruce, J. S. (1986). *The intuitive pragmatists: Conversations with chief executive officers*. Greensboro, NC: Center for Creative Leadership.
- Bryman, A. (1996). Leadership in organizations. In S. R. Clegg, C. Hardy, & W. R. Nord, *Handbook of organization studies* (pp. 276-292). Thousand Oaks, CA: Sage.
- Buchanan, G. M., & Seligman, M. E. P. (1995). Afterword: The future of the field. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 247-252). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Burns, J. M. (1978). *Leadership*. New York: Harper & Row.
- Bush, G. W. (2001a, September 20). Address to a joint session of Congress and the American people. Retrieved October 27, 2002, from <http://www.whitehouse.gov/news/releases/2001/09>
- Bush, G. W. (2001b, September 8). President emphasizes education reform in radio address. Retrieved November 26, 2002, from <http://www.whitehouse.gov/news/releases/2001/09>
- Bush, G. W. (2001c, September 1). President urges quick passage of education plan in radio address. Retrieved November 26, 2002, from <http://www.whitehouse.gov/news/releases/2001/08>
- Bush, G. W. (2001d, September 22). Radio address: Despite challenges, economy fundamentally strong. Retrieved November 26, 2002, from <http://www.whitehouse.gov/news/releases/2001/09>
- Bush, G. W. (2001e, August 18). Radio address of the president to the nation. Retrieved November 26, 2002, from <http://www.whitehouse.gov/news/releases/2001/08>
- Bush, G. W. (2001f, September 11). Statement by the president in his address to the nation. Retrieved October 27, 2002, from <http://www.whitehouse.gov/news>
- Bush, G. W. (2003). President Bush discusses top priorities for the U.S.: Press conference of the president. Retrieved November 29, 2003, from <http://www.whitehouse.gov/news>.
- Campbell, J. (2003). *Margaret Thatcher: Vol. 2. The iron lady*. London: Jonathan Cape.
- Chemers, M. M. (1997). *An integrative theory of leadership*. Mahwah, NJ: Lawrence Erlbaum.
- Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1987). A longitudinal study of moral judgment in U.S. males. In A. Colby & L. Kohlberg, *The measurement of moral judgment: Theoretical foundations and research validation* (Vol. 1, pp. 77-117). New York: Cambridge University Press.

- Colby, A., Kohlberg, L., & Kauffman, K. (1987). Theoretical introduction to the measurement of moral judgment. In A. Colby & L. Kohlberg, *The measurement of moral judgment: Theoretical foundations and research validation* (Vol. 1, pp. 1-61). New York: Cambridge University Press.
- Commons, M. L., Danaher, D., Miller, P. M., Goodheart, E. A., & Dawson, T. L. (2001). *Hierarchical complexity scoring system: How to score anything*. Unpublished manuscript.
- Commons, M. L., Krause, S. R., & Fayer, G. A. (1992). Institutional atmosphere, individual development, and the higher moral stages. Paper presented at the annual meeting of the Eastern Psychological Association, Boston, MA. (ERIC Document Reproduction Service No. ED348617)
- Commons, M. L., & Richards, F. A. (1984a). A general model of stage theory. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1: Late adolescent and adult cognitive development* (pp. 120-140). NY: Praeger.
- Commons, M. L., & Richards, F. A. (1984b). Applying the general stage model. In M. L. Commons, R. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1: Late adolescent and adult cognitive development* (pp. 141-157). NY: Praeger.
- Commons, M. L., & Richards, F. A. (2002). Organizing components into combinations: How stage transition works. *Journal of Adult Development*, 9(3), 159-177.
- Commons, M. L., & Richards, F. A. (2003). Four postformal stages. In J. Demick & C. Andreoletti (Eds.), *Handbook of adult development* (pp. 199-219). New York: Kluwer Academic/Plenum.
- Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review*, 18(3), 237-278.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Danaher, D. L., & Dawson, T. L. (1999, June). *An example of applying the general stage model to scoring the self section of the life experience interview*. Paper presented at the annual symposium of the Jean Piaget Society, Mexico City.
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4), 377-383.
- Dawson, T. L. (1999, June). *Scoring With the HCSS*. Paper presented at the annual symposium of the Jean Piaget Society, Mexico City.
- Dawson, T. L., Commons, M. L., & Wilson, M. (2001). *The shape of development*. Manuscript submitted for publication.
- Dawson, T. L., & Kay, A. (2001). *A stage is a stage is a stage: A direct comparison of two scoring systems*. Manuscript submitted for publication.
- Dawson, T. L., & Wilson, M. (in press). The LAAS: A computerized developmental scoring system for small- and large-scale assessments. *Educational Assessment*.

- Deison, P. V. (1994). Defining personal vision. *Dissertation Abstracts International*, 56(05), 1878A. (UMI No. 9530167)
- Dille, B. (2000). The prepared and spontaneous remarks of Presidents Regan and Bush: A validity comparison for at-a-distance measurements. *Political Psychology*, 21(3), 573-585.
- Dille, B., & Young, M. D. (2000). The conceptual complexity of presidents Carter and Clinton: An automated content analysis of temporal stability and source bias. *Political Psychology*, 21(3), 587-596.
- Dozier, J. B., & Miceli, M. P. (1985). Potential predictors of whistle-blowing: A prosocial behavior perspective. *Academy of Management Review*, 10(4), pp. 823-836.
- Ender, P. B. (2001). *Introduction to research design and statistics: Effect size*. Retrieved December 31, 2003, from University of California at Los Angeles Department of Education web site: <http://www.gseis.ucla.edu/courses/ed230a2/notes/effect.html>
- Fairhurst, G. T., & Sarr, R. A. (1996). *The art of framing: Managing the language of leadership*. San Francisco: Jossey-Bass.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(2), 477-531.
- Freud, S., & Bullitt, W. C. (1999). *Woodrow Wilson: A psychological study*. New Brunswick, NJ: Transaction Publishers.
- Frum, D. (2003). *The right man: The surprise presidency of George W. Bush*. New York: Random House.
- Gelderman, C. (1997). *All the presidents' words: The bully pulpit and the creation of the virtual presidency*. New York: Walker.
- George, J. M., & Bettenhausen, K. (1990). Understanding prosocial behavior, sales performance, and turnover: A group-level analysis in a service context. *Journal of Applied Psychology*, 75(6), 698-709.
- Gilligan, C. (1988a). Adolescent development reconsidered. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), *Mapping the moral domain* (pp. vii-xxxix). Cambridge, MA: Harvard University.
- Gilligan, C. (1988b). Preface. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), *Mapping the moral domain* (pp. i-v). Cambridge, MA: Harvard University.
- Gilligan, C. (1993). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Gilligan, C., & Attanucci, J. (1988). Two moral orientations. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), *Mapping the moral domain* (pp. 83-86). Cambridge, MA: Harvard University.
- Gould, S. J. (1989). *Wonderful life: The Burgess Shale and the nature of history*. New York: W. W. Norton.
- Hart, J. A. (1977). Cognitive maps of three Latin American policy makers. *World Politics*, 30(1), 115-140.

- Heenan, D. A., & Bennis, W. (1999). *Co-leaders: The power of great partnerships*. New York: Wiley.
- Hermann, M. G. (2003). Assessing leadership style: Trait analysis. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 178-212), Ann Arbor, MI: University of Michigan Press.
- Hersey, P., & Blanchard, K. H. (1969). Life cycle theory of leadership. *Training and Development Journal*, 23(5), 26-34.
- Hersey, P., & Blanchard, K. H. (1982). *Management of organizational behavior: Utilizing human resources* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hodgson, R. C., Levinson, D. J., & Zalesnik, A. (1965). *The executive role constellation: An analysis of personality and role relations in management*. Boston: Harvard University Division of Research.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.
- Janis, I. L. (1982). *Groupthink* (2nd ed.). Boston: Houghton Mifflin.
- Johnston, D. K. (1988). Adolescents' solutions to dilemmas in fables: Two moral orientations – Two problem solving strategies. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), *Mapping the moral domain* (pp. 49-71). Cambridge, MA: Harvard University.
- Kennan, N., & Hadley, M. (1986). The creation of political leaders in the context of American politics in the 1970s and 1980s. In C. F. Graumann & S. Moscovici (Eds.), *Changing conceptions of leadership* (pp. 145-169). New York: Springer Verlag.
- Kennedy, J. F. (1962a, October 22). Radio and television report to the American people on the Soviet arms buildup in Cuba. In Office of the Federal Register, *Public papers of the presidents of the United States: John F. Kennedy* (Vol. 1962, pp. 806-809). Washington, DC: U.S. Government Printing Office.
- Kennedy, J. F. (1962b, November 2). Radio and television remarks on the dismantling of Soviet missile bases in Cuba. In Office of the Federal Register, *Public papers of the presidents of the United States: John F. Kennedy* (Vol. 1962, p. 821). Washington, DC: U.S. Government Printing Office.
- Kohlberg, L. (1990). Which postformal levels are stages? In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. D. Sinnott (Eds.), *Adult Development, Vol. 2: Models and methods in the study of adolescent and adult thought* (pp. 263-268). New York: Praeger.
- Kohlberg, L., & Armon, C. (1984). Three types of stage models used in the study of adult development. In M. L. Commons, F. A. Richards, C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 383-394). New York: Praeger.
- Kouzes, J. M., & Posner, B. Z. (1987). *The leadership challenge: How to get extraordinary things done in organizations*. San Francisco: Jossey-Bass.

- Langer, W. C. (1972). *The mind of Adolf Hitler: The secret wartime report*. New York: Basic Books.
- Latané, B., & Darley, J. M. (1968). Group inhibition of bystander intervention in emergencies. *Journal of Personality and Social Psychology*, 10(3), 215-221.
- Lincoln, A. (1953). Message to Congress in special session. In R. P. Basler (Ed.), *The collected works of Abraham Lincoln* (Vol. 4, pp. 421-441). New Brunswick, NJ: Rutgers University Press. (Original work published July 4, 1861)
- Lincoln, A. (1953). Message to Congress. In R. P. Basler (Ed.), *The collected works of Abraham Lincoln* (Vol. 5, pp. 144-146). New Brunswick, NJ: Rutgers University Press. (Original work published March 6, 1862)
- Link, A. S. (Ed.). (1983). *The papers of Woodrow Wilson* (Vols. 41-42). Princeton, NJ: Princeton University Press.
- Lipman-Blumen, J. (1996). *The connective edge: Leading in an interdependent world*. San Francisco: Jossey-Bass.
- Locke, E. A. (2003). Leadership: Starting at the top. In C. L. Pearce & J. A. Conger (Eds.), *Shared leadership* (pp. 271-284). Thousand Oaks, CA: Sage.
- Lyons, N. P. (1982). *Conceptions of self and morality and modes of moral choice: Identifying justice and care in judgments of actual moral dilemmas*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- Lyons, N. P. (1983). Two perspectives: On self, relationships, and morality. *Harvard Educational Review*, 53(2), 125-144.
- Lyons, N. P. (1988). Two perspectives: On self, relationships, and morality. In C. Gilligan, J. V. Ward, & J. M. Taylor (Eds.), *Mapping the moral domain* (pp. 21-48). Cambridge, MA: Harvard University.
- Martin, J., & Frost, P. (1996). The organizational culture war games: A struggle for intellectual dominance. In S. R. Clegg, C. Hardy, & W. R. Nord (Eds.), *Handbook of organization studies* (pp. 599-621). Thousand Oaks, CA: Sage.
- Mayo, E. (1945). *The social problems of an industrial civilization*. Boston: Harvard University Graduate School of Business Administration.
- McDiarmid, J. (1937). Presidential inaugural addresses: A study in verbal symbols. *Public Opinion Quarterly*, 1(3), 79-82.
- Medhurst, M. J. (2003). Presidential speechwriting: Ten myths that plague modern scholarship. In K. Ritter & M. J. Medhurst (Eds.), *Presidential speechwriting: From the new deal to the Reagan revolution and beyond* (pp. 3-19). College Station, TX: Texas A&M University Press.
- Meir, G. (1972a, July 26). Statement to the Knesset. Retrieved March 31, 2003, from <http://www.mfa.gov.il>
- Meir, G. (1972b, October 16). Statement to the Knesset. Retrieved March 31, 2003, from <http://www.mfa.gov.il>
- Meir, G. (1973a, October 6). Broadcast to the nation [Radio broadcast]. Retrieved March 31, 2003, from <http://www.mfa.gov.il>

- Meir, G. (1973b, July 25). Statement to the Knesset [political section]. Retrieved March 31, 2003, from <http://www.mfa.gov.il>
- Meir, G. (1973c, October 16). Statement to the Knesset. Retrieved March 31, 2003, from <http://www.mfa.gov.il>
- Meir, G. (1973d, October 23). Statement to the Knesset. Retrieved March 31, 2003, from <http://www.mfa.gov.il>
- Milgram, S. (1969). The lost letter technique. *Psychology Today*, 3(1), 30-33, 66, 68-69.
- Nadler, D. A., & Tushman, M. L. (1989). Leadership for organizational change. In A. M. Mohrman Jr., S. A. Mohrman, G. E. Ledford Jr., T. G. Cummings, & E. E. Lawler III (Eds.), *Large-scale organizational change* (pp. 100-119). San Francisco: Jossey-Bass.
- Nanus, B. (1992). *Visionary leadership: Creating a compelling sense of direction for your organization*. San Francisco: Jossey-Bass.
- National Archives and Records Administration. (2001). Our heritage in documents: FDR's "day of infamy" speech: Crafting a call to arms. *Prologue: Quarterly of the National Archives and Records Administration*, 33(4). Retrieved March 30, 2002, from <http://www.nara.gov/publications/prologue/infamy.html>. Website changed to http://www.archives.gov/publications/prologue/winter_2001_crafting_day_of_infamy_speech.html
- National Archives and Records Administration. (2003). Joint address to Congress leading to a declaration of war against Japan (1941). Retrieved November 29, 2003, from <http://www.ourdocuments.gov>
- Neubert, M. J. (1999). Too much of a good thing or the more the merrier? Exploring the dispersion and gender composition of informal leadership in manufacturing teams. *Small Group Research*, 30(5), 635-646.
- Nystrom, P. C., & Starbuck, W. H. (1984). To avoid organizational crises, unlearn. *Organizational Dynamics*, 12(4), 53-65.
- O'Connor, J., & McDermott, I. (1997). *The art of systems thinking*. San Francisco, CA: Thorsons.
- Oettingen, G. (1995). Explanatory styles in the context of culture. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 209-224). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Office of the Federal Register. (1963). *Public papers of the presidents of the United States: John F. Kennedy* (Vol. 1962). Washington, DC: U.S. Government Printing Office.
- Office of the Federal Register. (1964). *Public papers of the presidents of the United States: John F. Kennedy* (Vol. 1963). Washington, DC: U.S. Government Printing Office.
- Oliver, C. R. (1993). Building the BEL system. *Security Management*, 37(5), 27-30.

- Oliver, C. R. (2003). *Briefing on pilot study for proposed dissertation titled impact of catastrophe on pivotal leaders' vision statements: Correspondences and discrepancies in moral reasoning, explanatory style, and rumination*. Unpublished manuscript, Fielding Graduate Institute, Santa Barbara, California.
- O'Toole, J., Galbraith, J., & Lawler, E. E., III. (2003). The promise and pitfalls of shared leadership: When two (or more) heads are better than one. In C. L. Pearce & J. A. Conger (Eds.), *Shared leadership* (pp. 250-267). Thousand Oaks, CA: Sage.
- Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, 32(4), 301-345.
- Pearce, C. L., & Conger, J. A. (Eds.). (2003a). *Shared leadership*. Thousand Oaks, CA: Sage.
- Pearce, C. L., & Conger, J. A. (2003b). All those years ago: The historical underpinnings of shared leadership. In C. L. Pearce & J. A. Conger, (Eds.), *Shared leadership* (pp. 1-18). Thousand Oaks, CA: Sage.
- Peterson, C. (1992). Explanatory style. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 376-382). New York: Cambridge University Press.
- Peterson, C. (2000). The future of optimism. *American Psychologist*, 55(1), 44-55.
- Peterson, C., Buchanan, G. M., & Seligman, M. E. P. (1995). Explanatory style: History and evolution of the field. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 1-20). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Peterson, C., Schulman, P., Castellon, C., & Seligman, M. E. P. (1992). The explanatory style scoring manual. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 383-400). New York: Cambridge University Press.
- Piaget, J. (1976). *The grasp of consciousness: Action and concept in the young child*. Cambridge, MA: Harvard University Press.
- Podell, J., & Anzovin, S. (2001). *Speeches of the American presidents* (2nd ed.). New York: H. W. Wilson.
- Porter, C. A., & Suedfeld, P. (1981). Integrative complexity in the correspondence of literary figures: Effects of personal and societal stress. *Journal of Personality and Social Psychology*, 40(2), 321-330.
- Post, J. M. (2003a). Leader personality assessments in support of government policy. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 39-61), Ann Arbor, MI: University of Michigan Press.
- Post, J. M. (Ed.). (2003b). *The psychological assessment of political leaders*, Ann Arbor, MI: University of Michigan Press.
- Raphael, T. D. (1982). Integrative complexity theory and forecasting international crises. *Journal of Conflict Resolution*, 26(3), 423-450.
- Reivich, K. (1995). The measurement of explanatory style. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 21-47). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Rettew, D., & Reivich, K. (1995). Sports and explanatory style. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 173-185). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Richards, F. A., & Commons, M. L. (1984). Systematic, metasystematic, and cross-paradigmatic reasoning: A case for stages of reasoning beyond formal operations. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 92-119). New York: Praeger.
- Ritter, K., & Medhurst, M. J. (Eds.). (2003). *Presidential speechwriting: From the new deal to the Reagan revolution and beyond*. College Station, TX: Texas A&M University Press.
- Roberts, C. W. (1997). Introduction. In C. W. Roberts (Ed.), *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts* (pp. 1-8). Mahwah, NJ: Lawrence Erlbaum Associates.
- Robson, C. (1993). *Real world research: A resource for social scientists and practitioner-researchers*. Malden, MA: Blackwell.
- Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker: An account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago*. Cambridge, MA: Harvard University Press.
- Roosevelt, F. D. (1950). Address to the Congress asking that a state of war be declared between the United States and Japan. In S. I. Rosenman (Ed.), *The public papers and addresses of Franklin D. Roosevelt: 1941 vol.* (pp. 514-516). New York: Harper & Brothers. (Original work published December 8, 1941)
- Rosenman, S. I. (Ed.). (1950). *The public papers and addresses of Franklin D. Roosevelt* (Vol. 1941). New York: Harper & Brothers.
- Satterfield, J. M. (1998). Cognitive-affective states predict military and political aggression and risk taking. *Journal of Conflict Resolution*, 42(6), 667-690.
- Satterfield, J. M., & Seligman, M. E. P. (1994). Military aggression and risk predicted by explanatory style. *Psychological Science*, 5(2), 77-82.
- Schafer, M. (2000). Issues in assessing psychological characteristics at a distance: An introduction to the symposium. *Political Psychology*, 21(3), 511-527.
- Schein, E. H. (1985). *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Schein, E. H. (1992). *Organizational culture and leadership* (2nd ed.). San Francisco: Jossey-Bass.
- Schulman, P., Castellon, C., & Seligman, M. E. P. (1989). Assessing explanatory style: The content analysis of verbatim explanations and the attributional style questionnaire. *Behaviour Research and Therapy*, 27(5), 505-512.
- Seibert, S. C., Sparrowe, R. T., & Liden, R. C. (2003). A group exchange structure approach to leadership in groups. In C. L. Pearce & J. A. Conger (Eds.), *Shared leadership* (pp. 173-192). Thousand Oaks, CA: Sage.

- Seligman, M. E. P. (1998). *Learned optimism: How to change your mind and your life*. New York: Pocket Books.
- Shakespeare, W. (1997). Othello. In G. B. Evans, *The Riverside Shakespeare (2nd ed.): The complete works* (pp. 1246-1296). Boston: Houghton Mifflin. (Original work published 1604)
- Shakespeare, W. (1997). The tempest. In G. B. Evans, *The Riverside Shakespeare (2nd ed.): The complete works* (pp. 1656-1688). Boston: Houghton Mifflin. (Original work published 1611)
- Shapiro, G. (1997). The future of coders: Human judgments in a world of sophisticated software. In C. W. Roberts, *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts* (pp. 225-238). Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, C. P. (1992a). Introduction: Inferences from verbal material. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 1-17). New York: Cambridge University Press.
- Smith, C. P. (1992b). Reliability issues. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 515-536). New York: Cambridge University Press.
- Smith, C. P., Feld, S. C., Franz, C. E. (1992). Methodological considerations: Steps in research employing content analysis systems. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 515-536). New York: Cambridge University Press.
- Snyder, C. R., & Lopez, S. J. (Eds.). (2002). *Handbook of positive psychology*. New York: Oxford University Press.
- SPSS. (1999). *SPSS base 10.0 applications guide*. Chicago, IL: SPSS, Inc.
- Stevens-Long, J., & Commons, M. L. (1992). *Adult life: Developmental processes* (4th ed.). Mountain View, CA: Mayfield.
- Stogdill, R. M. (1974). *Handbook of leadership: A survey of theory and research*. New York: The Free Press.
- Suedfeld, P., & Bluck, S. (1988). Changes in integrative complexity prior to surprise attacks. *Journal of Conflict Resolution*, 32(4), 626-635.
- Suedfeld, P., Guttieri, K., & Tetlock, P. E. (2003). Assessing integrative complexity at a distance: Archival analyses of thinking and decision making. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 246-270). Ann Arbor, MI: University of Michigan Press.
- Suedfeld, P., & Tetlock, P. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolution*, 21(1), 169-184.
- Suedfeld, P., Tetlock, P. E., & Ramirez, C. (1977). War, peace, and integrative complexity: U.N. speeches on the Middle East problem, 1947-1976. *Journal of Conflict Resolution*, 21(3), 427-440.

- Tappan, M., Kohlberg, L., Schrader, D., Higgins, A., Armon, C., & Lei, T. (1987). Heteronomy and autonomy in moral development: Two types of moral judgments. In A. Colby & L. Kohlberg, *The measurement of moral judgment: Theoretical foundations and research validation* (Vol. 1, pp. 315-380). New York: Cambridge University Press.
- Tetlock, P. E. (1979). Identifying victims of groupthink from public statements of decision makers. *Journal of Personality and Social Psychology*, 37(8), 1314-1324.
- Tetlock, P. E. (1981). Pre- to post election shifts in Presidential rhetoric: Impression management or cognitive adjustment? *Journal of Personality and Social Psychology*, 41(2), 207-212.
- Thatcher, M. (1981a, November 4). House of Commons speech: Debate on address. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Thatcher, M. (1981b, July 27). House of Commons speech: Economic censure. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Thatcher, M. (1981c, November 28). House of Commons speech: Economic censure. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Thatcher, M. (1982a, April 3). House of Commons speech: Falkland Islands. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Thatcher, M. (1982b, April 14). House of Commons speech: Falkland Islands. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Thatcher, M. (1982c, May 20). House of Commons speech: Falkland Islands. Retrieved November 27, 2002, from <http://www.margaretthatcher.org>
- Turner, N., Barling, J., & Zacharatos, A. (2002). Positive psychology at work. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 715-728). New York: Oxford University Press.
- UCLA Academic Technology Services (2002). *SPSS FAQ*. Retrieved December 27, 2002, from <http://www.ats.ucla.edu/stat/spss/faq/alpha.html>
- Walker, S. G., & Schafer, M. (2000). The political universe of Lyndon B. Johnson and his advisors: Diagnostic and strategic propensities in their operational codes. *Political Psychology*, 21(3), 529-543.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Non-reactive research in the social sciences*. Chicago: Rand McNally.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (2000). *Unobtrusive measures* (Rev. ed.). Thousand Oaks, CA: Sage.
- Weick, K. E. (2001). *Making sense of the organization*. Malden, MA: Blackwell.
- Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco: Jossey-Bass.
- Wilson, W. (1983). Address to a joint session of Congress. In A. S. Link (Ed.), *The papers of Woodrow Wilson* (Vol. 41, pp. 519-527). Princeton, NJ: Princeton University Press. (Original work published April 2, 1917)

- Windt, Jr., T. O. (2003). John F. Kennedy: Presidential speechwriting as rhetorical collaboration. In K. Ritter & M. J. Medhurst (Eds.), *Presidential speechwriting: From the new deal to the Reagan revolution and beyond*. College Station, TX: Texas A&M University Press.
- Winter, D. G. (1991). Measuring personality at a distance: Development of an integrated system for scoring motives in running text. In A. J. Stewart, J. M. Healy Jr., D. Ozer, & R. Hogan (Eds.), *Perspectives in personality, A research annual: Approaches to understanding lives* (Vol. 3B, pp. 59-89). London: Jessica Kingsley.
- Winter, D. G. (1992). Content analysis of archival materials, personal documents, and everyday verbal productions. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 110-125). New York: Cambridge University Press.
- Winter, D. G. (2002). Motivation and political leadership. In L. O. Valenty & O. Feldman (Eds.), *Political leadership for the new century: Personality and behavior among American leaders* (pp. 25-47). Westport, CT: Praeger.
- Winter, D. G. (2003a). Assessing leaders' personalities: A historical survey of academic research studies. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 11-38). Ann Arbor, MI: University of Michigan Press.
- Winter, D. G. (2003b). Measuring the motives of political actors at a distance. In J. M. Post (Ed.), *The psychological assessment of political leaders* (pp. 153-177). Ann Arbor, MI: University of Michigan Press.
- Winter, D. G., & Stewart, A. J. (1977). Content analysis as a technique for assessing political leaders. In M. G. Hermann & T. W. Milburn (Eds.), *A psychological examination of political leaders* (pp. 27-61). New York: Free Press.
- Yu, A. (n.d.). *Using SAS for item analysis and test construction*. Retrieved December 28, 2002, from <http://seamonkey.ed.asu.edu/~alex/teaching/assessment/alpha.html>
- Zullov, H. M. (1988). Pessimistic rumination predicts defeat of presidential candidates: 1900-1988. *Dissertation Abstracts International* 49(09), 4070B (UMI No. 8824814)
- Zullov, H. M. (1995). Pessimistic rumination in American politics and society. In G. M. Buchanan & M. E. P. Seligman (Eds.), *Explanatory style* (pp. 187-208). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Zullov, H. M., Oettingen, G., Peterson, C., & Seligman, M. E. P. (1988). Pessimistic explanatory style in the historical record: CAVing LBJ, presidential candidates, and East versus West Berlin. *American Psychologist*, 43(9), 673-682.
- Zullov, H. M., & Seligman, M. E. P. (1990). Pessimistic rumination predicts defeat of presidential candidates, 1900 to 1984. *Psychological Inquiry*, 1(1), 52-61.

Appendix A
Scoring Aid for Moral Reasoning Stage

Stage No.	Vocalizations and Words	Verbal Relationships	Generic Examples	Vision Examples	Discriminations
0	None				Exact – no generalization
1	Babble				Rote, generalized
2	Phonemes		Gestures, "aa"		Open-ended classes
3	Morphemes	Phoneme action relationships	Holding cup, saying "wa"		Concepts
4	Single words	Concepts related through a word	Says "water"		Relations among concepts
5	Pronouns	Incomplete sentences, complete sentences, crude dichotomies, ordered numbers and letters	States rule: "Don't touch"		Imitates and acquires sequences. Follows short sequential acts.
6	Connectives: as, when, then, why, before	Multiple sentence stories with logical order, sequenced events	Says "Water is filling the bathtub."	The committee – the civil contingencies committee – will meet again tomorrow at 8 am.	Simple deduction but contradiction not excluded. Follows lists of sequential acts.

Appendix A (continued)

Stage No.	Vocalizations and Words	Verbal Relationships	Generic Examples	Vision Examples	Discriminations
7	Times, places, acts, actors	Stories with events logically ordered, sequenced and connected to reality with some specified times, places, acts, actors	"Water is filling the bathtub so I can take a bath."	Private flights have been stopped except where specifically authorized.	Simple logical deduction and empirical rules involving time sequence. Simple arithmetic.
8	Interactions, social events, what happened among others	Stories about things, incidents, events, actors, actions, places in context of self and other	"If I turn on the faucet, the water will run out and fill the bathtub. My parents like it when I fill the tub myself. They are more willing to buy me clothes when I am always clean."	I therefore spoke to President Roosevelt on the Atlantic telephone last night with a view to arranging the timing of our respective declarations.	Full arithmetic

Appendix A (continued)

Stage No.	Vocalizations and Words	Verbal Relationships	Generic Examples	Vision Examples	Discriminations
9	Variable time, place, act, actor, state, type; "All, None, Some"	Variable time, place, act, actor, state, type. Stereotypes, "all, none, some, never, sometimes, always, no one, somebody, everyone."	"Water can run out quickly or slowly and the temperature may get colder after a while. All teachers do that!"	Every circumstance of calculated and characteristic Japanese treachery was employed.	Discriminates variables, e.g., stereotypes, all, none, some
10	Linear, logical, one dimensional	Variables connected by "If...then" and other relations. Empirical or logical evidence and argument. Relations among variables.	"Turning a faucet changes the size of the opening of the passage way and thereby controls the amount of flow."	If you are attacked, [then] defend yourselves.	Argue using empirical or logical evidence. Logic is linear, one dimensional.

Appendix A (continued)

Stage No.	Vocalizations and Words	Verbal Relationships	Generic Examples	Vision Examples	Discriminations
11	Systems words: legal system, society, our company, the economy, the country	Systems of relations. Relationships among variables.	"For slow flows, amount of flow is controlled by 1) water pressure, 2) size of faucet opening, 3) smallest of pipe sizes to source."	It will involve the organization and mobilization of all the material resources of the country to supply the materials of war and serve the incidental needs of the nation in the most abundant and yet the most economical and efficient way possible.	Constructs multi-variate systems and matrices, coordinating more than one variable. Events and ideas situated in a larger context.
12	Metalogical, meta-analytic words. Properties of systems.	Super-systems of systems	"Systems that describe non-turbulent water flow and electric flow (amperes) are isomorphic. Flow equals pressure (volts) x resistance (ohms)."	It only remains for the two great democracies to face their task with whatever strength God may give them, now that the issue is joined in a most direct manner.	Integrate systems to construct multi-systems. Compare systems and perspectives in systematic way across multiple domains. Reflect on systems.

Appendix A (continued)

Stage No.	Vocalizations and Words	Verbal Relationships	Generic Examples	Vision Examples	Discriminations
13	Paradigmatic words		"There are many metasystems that belong to non-relativistic and non-statistical mechanics."		Discriminate how to fit metasystems together to form paradigms
14	Integration of paradigms		"By crossing mathematics and physical experimentation, classical mechanics arises."		Discriminate how to form new fields by crossing paradigms.

Note: Adapted from *Hierarchical complexity scoring system: How to score anything*, by M. L. Commons, D. Danaher, P. M. Miller, E. A. Goodheart, and T. L. Dawson, 2001. Unpublished manuscript. Adapted with permission.

Appendix B

Scoring Aid for Moral Reasoning Orientation

General instruction: Assess the speaker's worldview expressed by the text. Score the text as follows:

Only Caring when the text expresses only a caring worldview.

Mostly Caring when the text expresses both the caring and justice worldviews but the caring worldview predominates.

Mostly Justice when the text expresses both the caring and justice worldviews but the justice worldview predominates.

Only Justice when the text expresses only a justice worldview.

- "Caring" is a worldview based on human interrelationships. Typical content includes one or more of the following:

- Alleviate burden or suffering
- Foster welfare of another
- Maintain or restore interpersonal relationships
- Avoid interpersonal conflict
- Interdependence
- Emphasis on situation over principle
- Effects on others

- "Justice" is a worldview based on laws, rules and fairness. Typical content includes one or more of the following:

- Laws
- Standards / rules / principles
- Obligation / duty / commitment
- Fairness (how I would like to be treated if in the other's place)
- Others have their own contexts
- Emphasis on principle over situation
- Effect on self (includes "trouble" and "how to decide")

Appendix C
Scoring Aid for Explanatory Style

Scale Point	Internality Versus Externality	Stability Versus Instability of <i>Cause</i> (Not Event)	Globality Versus Specificity
1	External. Speaker attributes blame or credit to someone or something completely external to self or "my country" or "my organization."	Unstable. Cause goes away: it will never be present again.	Specific. Cause affects one area of functioning (e.g., either achievement or affiliation), or one area of public policy, or one specific group of people.
2	Mixed (more external). Speaker attributes cause to combination of internal and external factors.	Mixed (more unstable). Cause cannot occur again but may have ongoing influence.	Mixed (more specific). Cause affects part of one category (e.g., achievement or affiliation) and perhaps part of another.
3	Mixed (somewhat more external). Speaker attributes cause to combination of internal and external factors.	Mixed (somewhat unstable). Cause likely to occur again intermittently, or may affect behavior intermittently.	Mixed (somewhat specific). Cause affects more of one category (e.g., achievement or affiliation) and possibly more of another.
4	Mixed (external and internal). Speaker attributes cause to combination of internal and external factors.	Mixed. Cause likely to occur again intermittently and may have ongoing influence on behavior.	Mixed. Cause affects parts of several categories (e.g., achievement and affiliation).
5	Mixed (somewhat more internal). Speaker attributes cause to combination of internal and external factors.	Mixed (somewhat stable). Cause likely to occur again, or does exert ongoing influence on behavior.	Mixed (somewhat global). Cause affects more of more categories (e.g., achievement and affiliation).

Appendix C (Continued)

Scale Point	Internality Versus Externality	Stability Versus Instability of <i>Cause</i> (Not Event)	Globality Versus Specificity
6	Mixed (more internal). Speaker attributes cause to combination of internal and external factors.	Mixed (more stable). Cause probably will occur again and is characterological.	Mixed (more global). Cause affects still more of more categories (e.g., achievement and affiliation).
7	Internal. Speaker attributes cause to any characteristic solely internal to self or "my country" or "my organization."	Stable. Cause will persist long into the future: it is unalterable and continuous.	Global. Cause affects many areas of functioning (e.g., virtually all of achievement and affiliation), or many areas of public policy, or most people in a very large group.

Note: If information is insufficient to assign a rating, assign a 4.

Appendix D

Scoring Aid for Rumination

Rate each sentence for presence of rumination. Score rumination present if, from the speaker's point of view, the sentence:

- focuses on a bad event—the who, what, where, and how of a bad event, or
- explains a bad event—the why, or
- presents a negative emotion—"a word is used to indicate the presence of an emotional state or emotional display (e.g., sad, depressed, angry, crying)"

Appendix E
Stages of Moral Reasoning Displayed Numerically

Leader	Vision	Expert's Score	Stage of Moral Reasoning
Lincoln	1	16	Dawson: 1st transition to abstract systems Commons: Formal with some systematic
Lincoln	2	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Lincoln	3	18	Dawson: Consolidated abstract systems Commons: Systematic
Lincoln	4	19	Dawson: 1st transition to single principles Commons: Systematic with some metasytematic
Lincoln	5	18	Dawson: Consolidated abstract systems Commons: Systematic
Lincoln	6	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Wilson	1	18	Dawson: Consolidated abstract systems Commons: Systematic
Wilson	2	19	Dawson: 1st transition to single principles Commons: Systematic with some metasytematic
Wilson	3	20	Dawson: 2nd transition to single principles Commons: Metasytematic with some systematic
Wilson	4	20	Dawson: 2nd transition to single principles Commons: Metasytematic with some systematic
Wilson	5	19	Dawson: 1st transition to single principles Commons: Systematic with some metasytematic
Wilson	6	20	Dawson: 2nd transition to single principles Commons: Metasytematic with some systematic
Roosevelt	1	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Roosevelt	2	16	Dawson: 1st transition to abstract systems Commons: Formal with some systematic
Roosevelt	3	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal

Appendix E (Continued)

Leader	Vision		Expert's Score: Stage of Moral Reasoning
Roosevelt	4	15	Dawson: Consolidated abstract mappings Commons: Formal
Roosevelt	5	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Roosevelt	6	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Kennedy	1	18	Dawson: Consolidated abstract systems Commons: Systematic
Kennedy	2	18	Dawson: Consolidated abstract systems Commons: Systematic
Kennedy	3	19	Dawson: 1st transition to single principles Commons: Systematic with some metasystematic
Kennedy	4	18	Dawson: Consolidated abstract systems Commons: Systematic
Kennedy	5	—	Unscorable
Kennedy	6	18	Dawson: Consolidated abstract systems Commons: Systematic
Meir	1	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Meir	2	18	Dawson: Consolidated abstract systems Commons: Systematic
Meir	3	18	Dawson: Consolidated abstract systems Commons: Systematic
Meir	4	15	Dawson: Consolidated abstract mappings Commons: Formal
Meir	5	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Meir	6	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal

Appendix E (Continued)

Leader	Vision		Expert's Score: Stage of Moral Reasoning
Thatcher	1	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Thatcher	2	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Thatcher	3	18	Dawson: Consolidated abstract systems Commons: Systematic
Thatcher	4	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Thatcher	5	18	Dawson: Consolidated abstract systems Commons: Systematic
Thatcher	6	18	Dawson: Consolidated abstract systems Commons: Systematic
Bush	1	16	Dawson: 1st transition to abstract systems Commons: Formal with some systematic
Bush	2	16	Dawson: 1st transition to abstract systems Commons: Formal with some systematic
Bush	3	15	Dawson: Consolidated abstract mappings Commons: Formal
Bush	4	15	Dawson: Consolidated abstract mappings Commons: Formal
Bush	5	17	Dawson: 2nd transition to abstract systems Commons: Systematic with some formal
Bush	6	16	Dawson: 1st transition to abstract systems Commons: Formal with some systematic

Appendix F

Unweighted Moral Reasoning Orientation Ratings Displayed Numerically

Leader	Vision	Only Caring	Mostly Caring	Mostly Justice	Only Justice
Lincoln	1	0.00%	66.67%	33.33%	0.00%
Lincoln	2	0.00%	37.50%	25.00%	37.50%
Lincoln	3	10.53%	31.58%	10.53%	47.37%
Lincoln	4	4.00%	24.00%	32.00%	40.00%
Lincoln	5	4.76%	19.05%	23.81%	52.38%
Lincoln	6	0.00%	16.67%	83.33%	0.00%
Wilson	1	0.00%	0.00%	33.33%	66.67%
Wilson	2	0.00%	0.00%	33.33%	66.67%
Wilson	3	0.00%	33.33%	66.67%	0.00%
Wilson	4	0.00%	22.22%	44.44%	33.33%
Wilson	5	0.00%	33.33%	33.33%	33.33%
Wilson	6	33.33%	66.67%	0.00%	0.00%
Roosevelt	1	0.00%	16.67%	16.67%	66.67%
Roosevelt	2	0.00%	66.67%	0.00%	33.33%
Roosevelt	3	0.00%	50.00%	50.00%	0.00%
Roosevelt	4	0.00%	0.00%	33.33%	66.67%
Roosevelt	5	22.22%	11.11%	33.33%	33.33%
Roosevelt	6	0.00%	66.67%	33.33%	0.00%
Kennedy	1	0.00%	12.50%	75.00%	12.50%
Kennedy	2	16.70%	0.00%	33.33%	50.00%
Kennedy	3	0.00%	50.00%	33.33%	16.67%
Kennedy	4	0.00%	0.00%	66.67%	33.33%
Kennedy	5			Unscorable	
Kennedy	6	0.00%	11.11%	22.22%	66.67%
Meir	1	0.00%	20.00%	40.00%	40.00%
Meir	2	0.00%	16.67%	50.00%	33.33%
Meir	3	0.00%	0.00%	100.00%	0.00%
Meir	4			Unscorable	
Meir	5	0.00%	0.00%	66.67%	33.33%
Meir	6	11.11%	22.22%	11.11%	55.56%
Thatcher	1	0.00%	31.25%	37.50%	31.25%
Thatcher	2	0.00%	15.38%	42.31%	42.31%
Thatcher	3	0.00%	0.00%	20.00%	80.00%
Thatcher	4	0.00%	7.14%	57.14%	35.71%
Thatcher	5	0.00%	8.33%	50.00%	41.67%
Thatcher	6	0.00%	25.00%	16.67%	58.33%

Appendix F (Continued)

Leader	Vision	Only Caring	Mostly Caring	Mostly Justice	Only Justice
Bush	1	0.00%	100.00%	0.00%	0.00%
Bush	2	20.00%	60.00%	20.00%	0.00%
Bush	3	16.67%	66.67%	16.67%	0.00%
Bush	4	16.67%	50.00%	16.67%	16.67%
Bush	5	5.00%	25.00%	35.00%	35.00%
Bush	6	0.00%	0.00%	100.00%	0.00%

Appendix G
Explanatory Style Ratings Displayed Numerically

Leader	Vision	Bad/Good	Internality/ Externality	Stability/ Instability	Globality/ Specificity	Total
Lincoln	1	Bad		No Extractions		
Lincoln	1	Good	4.32	4.83	3.33	12.48
Lincoln	2	Bad	4.29	4.21	4.33	12.83
Lincoln	2	Good	5.22	3.56	3.11	11.89
Lincoln	3	Bad	4.10	3.20	3.13	10.43
Lincoln	3	Good	4.13	4.33	3.67	12.13
Lincoln	4	Bad	3.10	3.18	3.33	9.61
Lincoln	4	Good	4.57	4.80	4.27	13.63
Lincoln	5	Bad	3.50	3.63	3.27	10.40
Lincoln	5	Good	4.39	4.14	3.58	12.11
Lincoln	6	Bad	5.50	3.00	2.50	11.00
Lincoln	6	Good		No Extractions		
Wilson	1	Bad	2.67	4.50	4.25	11.42
Wilson	1	Good		No Extractions		
Wilson	2	Bad	3.17	3.97	3.37	10.50
Wilson	2	Good		No Extractions		
Wilson	3	Bad	2.25	3.78	4.79	10.82
Wilson	3	Good	4.15	3.65	4.26	12.07
Wilson	4	Bad		No Extractions		
Wilson	4	Good		No Extractions		
Wilson	5	Bad	2.33	4.33	4.67	11.33
Wilson	5	Good		No Extractions		
Wilson	6	Bad	1.87	3.36	5.25	10.48
Wilson	6	Good	6.00	5.67	3.17	14.83
Roosevelt	1	Bad	2.46	3.36	4.61	10.43
Roosevelt	1	Good		No Extractions		
Roosevelt	2	Bad		No Extractions		
Roosevelt	2	Good	5.33	2.67	3.67	11.67
Roosevelt	3	Bad	2.28	3.22	4.64	10.14
Roosevelt	3	Good	7.00	7.00	4.00	18.00
Roosevelt	4	Bad	2.42	3.45	4.81	10.67
Roosevelt	4	Good	1.50	2.50	7.00	11.00
Roosevelt	5	Bad		No Extractions		
Roosevelt	5	Good		No Extractions		
Roosevelt	6	Bad		No Extractions		

Appendix G (Continued)

Leader	Vision	Bad/Good	Internality/ Externality	Stability/ Instability	Globality/ Specificity	Total
Roosevelt	6	Good		No Extractions		
Kennedy	1	Bad	3.56	4.15	4.11	11.82
Kennedy	1	Good	5.50	5.50	7.00	18.00
Kennedy	2	Bad	3.72	3.31	3.36	10.39
Kennedy	2	Good	5.06	4.30	3.72	13.07
Kennedy	3	Bad	3.08	3.38	2.71	9.17
Kennedy	3	Good	2.87	3.83	2.70	9.40
Kennedy	4	Bad	2.42	3.80	4.95	11.18
Kennedy	4	Good		No Extractions		
Kennedy	5	Bad		No Extractions		
Kennedy	5	Good	3.00	3.67	5.33	12.00
Kennedy	6	Bad	3.29	3.79	3.79	10.88
Kennedy	6	Good	4.38	4.78	4.31	13.46
Meir	1	Bad	1.25	3.21	5.50	9.96
Meir	1	Good	3.83	4.33	4.88	13.04
Meir	2	Bad	1.61	4.09	5.30	11.00
Meir	2	Good	3.00	4.33	4.00	11.33
Meir	3	Bad		No Extractions		
Meir	3	Good		No Extractions		
Meir	4	Bad	2.44	3.69	4.24	10.37
Meir	4	Good	4.50	4.19	4.11	12.81
Meir	5	Bad		No Extractions		
Meir	5	Good		No Extractions		
Meir	6	Bad	2.33	5.00	5.00	12.33
Meir	6	Good	5.50	4.00	5.50	15.00
Thatcher	1	Bad	2.83	4.03	3.60	10.47
Thatcher	1	Good	4.01	4.68	3.12	11.82
Thatcher	2	Bad	3.75	3.75	4.39	11.89
Thatcher	2	Good	4.67	4.83	4.08	13.58
Thatcher	3	Bad	3.17	4.50	5.67	13.33
Thatcher	3	Good		No Extractions		
Thatcher	4	Bad	3.13	3.58	4.23	10.94
Thatcher	4	Good	2.33	4.06	3.67	10.06
Thatcher	5	Bad		No Extractions		
Thatcher	5	Good		No Extractions		
Thatcher	6	Bad	3.00	4.44	4.56	12.00
Thatcher	6	Good		No Extractions		

Appendix G (Continued)

Leader	Vision	Bad/Good	Internality/ Externality	Stability/ Instability	Globality/ Specificity	Total
Bush	1	Bad	3.43	3.47	3.73	10.63
Bush	1	Good	4.20	4.00	4.27	12.47
Bush	2	Bad		No Extractions		
Bush	2	Good	5.92	5.92	3.75	15.58
Bush	3	Bad	2.50	4.00	2.50	9.00
Bush	3	Good	5.42	4.42	3.92	13.75
Bush	4	Bad	2.50	4.17	5.20	11.87
Bush	4	Good		No Extractions		
Bush	5	Bad	2.57	4.33	5.57	12.48
Bush	5	Good	3.28	4.11	5.06	12.44
Bush	6	Bad		No Extractions		
Bush	6	Good	3.50	5.00	6.00	14.50

Appendix H
Rumination Ratings Displayed Numerically

Leader	Vision	Rating	No. of Sentences
Lincoln	1	13.89%	36
Lincoln	2	23.26%	43
Lincoln	3	30.66%	137
Lincoln	4	29.82%	218
Lincoln	5	28.90%	218
Lincoln	6	36.84%	19
Wilson	1	55.17%	29
Wilson	2	39.58%	48
Wilson	3	26.67%	60
Wilson	4	50.98%	102
Wilson	5	31.25%	48
Wilson	6	49.06%	106
Roosevelt	1	65.25%	141
Roosevelt	2	33.33%	21
Roosevelt	3	61.21%	116
Roosevelt	4	76.92%	26
Roosevelt	5	57.86%	140
Roosevelt	6	55.56%	27
Kennedy	1	46.21%	132
Kennedy	2	24.69%	162
Kennedy	3	29.27%	41
Kennedy	4	69.51%	82
Kennedy	5	44.44%	9
Kennedy	6	36.07%	219
Meir	1	44.17%	206
Meir	2	62.50%	72
Meir	3	42.65%	136
Meir	4	75.00%	32
Meir	5	61.22%	147
Meir	6	45.63%	103
Thatcher	1	17.37%	259
Thatcher	2	20.66%	305
Thatcher	3	21.47%	177
Thatcher	4	38.10%	147
Thatcher	5	29.46%	112
Thatcher	6	28.33%	180

Appendix H (continued)

Leader	Vision	Rating	No. of Sentences
Bush	1	27.27%	33
Bush	2	3.23%	31
Bush	3	0.00%	32
Bush	4	57.14%	35
Bush	5	60.22%	181
Bush	6	28.13%	32

Appendix I
Table of 42 Selected Vision Statements

Identifier	Vision Statement	Source
Lincoln 1	Speech at Cincinnati, Ohio, February 12, 1861	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 4, pp. 197-200). New Brunswick, NJ: Rutgers University Press.
Lincoln 2	Speech at Pittsburgh, Pennsylvania, February 15, 1861	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 4, pp. 210-214). New Brunswick, NJ: Rutgers University Press.
Lincoln 3	First inaugural address, March 4, 1861	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 4, pp. 262-271). New Brunswick, NJ: Rutgers University Press.
Lincoln 4	Message to Congress in special session, July 4, 1861	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 4, pp. 421-441). New Brunswick, NJ: Rutgers University Press.
Lincoln 5	Annual message to Congress, December 3, 1861	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 5, pp. 35-53). New Brunswick, NJ: Rutgers University Press.
Lincoln 6	Message to Congress, March 6, 1862	Basler, R. P. (Ed.). (1953). <i>The Collected Works of Abraham Lincoln</i> (Vol. 5, pp. 144-146). New Brunswick, NJ: Rutgers University Press.
Wilson 1	An address to a joint session of Congress, February 3, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 41, pp. 108-112). Princeton, NJ: Princeton University Press.

Appendix I (Continued)

Identifier	Vision Statement	Source
Wilson 2	An address to a joint session of Congress, February 26, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 41, pp. 283-287). Princeton, NJ: Princeton University Press.
Wilson 3	The second inaugural address, March 5, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 41, pp. 332-336). Princeton, NJ: Princeton University Press.
Wilson 4	An address to a joint session of Congress, April 2, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 41, pp. 519-527). Princeton, NJ: Princeton University Press.
Wilson 5	An appeal to the American people, April 15, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 42, pp. 71-75). Princeton, NJ: Princeton University Press.
Wilson 6	A Flag Day address, June 14, 1917	Link, A. S. (Ed.). (1983). <i>The papers of Woodrow Wilson</i> (Vol. 42, pp. 498-504). Princeton, NJ: Princeton University Press.
Roosevelt 1	Fireside chat, September 11, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 384-392). New York: Harper & Brothers.
Roosevelt 2	Radio address on community mobilization for human needs, October 3, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 404-405). New York: Harper & Brothers.
Roosevelt 3	Address for Navy and Total Defense Day, October 27, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 438-444). New York: Harper & Brothers.

Appendix I (Continued)

Identifier	Vision Statement	Source
Roosevelt 4	Address to Congress asking that a state of war be declared between the United States and Japan, December 8, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 514-516). New York: Harper & Brothers.
Roosevelt 5	Fireside chat, December 9, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 522-531). New York: Harper & Brothers.
Roosevelt 6	Radio address on the 150th anniversary of the ratification of the Bill of Rights, December 15, 1941	Rosenman, S. I. (Ed.). (1950). <i>The public papers and addresses of Franklin D. Roosevelt</i> (Vol. 1941, pp. 554-557). New York: Harper & Brothers.
Kennedy 1	Radio and television address to the American people: Nuclear testing and disarmament, March 2, 1962	Office of the Federal Register. (1963). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1962, pp. 186-192). Washington, DC: U.S. Government Printing Office.
Kennedy 2	Radio and television report to the American people on the state of the national economy, August 13, 1962	Office of the Federal Register. (1963). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1962, pp. 611-617). Washington, DC: U.S. Government Printing Office.
Kennedy 3	Radio and television report to the nation on the situation at the University of Mississippi, September 30, 1962	Office of the Federal Register. (1963). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1962, pp. 726-728). Washington, DC: U.S. Government Printing Office.

Appendix I (Continued)

Identifier	Vision Statement	Source
Kennedy 4	Radio and television report to the American people on the Soviet arms buildup in Cuba, October 22, 1962	Office of the Federal Register. (1963). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1962, pp. 806-809). Washington, DC: U.S. Government Printing Office.
Kennedy 5	Radio and television remarks on the dismantling of Soviet missile bases in Cuba, November 2, 1962	Office of the Federal Register. (1963). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1962, p. 821). Washington, DC: U.S. Government Printing Office.
Kennedy 6	Annual message to the Congress on the state of the union, January 14, 1963	Office of the Federal Register. (1964). <i>Public papers of the presidents of the United States: John F. Kennedy</i> (Vol. 1963, pp. 11-19). Washington, DC: U.S. Government Printing Office.
Meir 1	Statement to the Knesset by Prime Minister Meir, July 26, 1972	Retrieved March 31, 2003, from http://www.mfa.gov.il
Meir 2	Statement to the Knesset by Prime Minister Meir, October 16, 1972	Retrieved March 31, 2003, from http://www.mfa.gov.il
Meir 3	Statement to the Knesset by the prime minister [political section], July 25, 1973	Retrieved March 31, 2003, from http://www.mfa.gov.il
Meir 4	Broadcast to the nation by Prime Minister Meir, October 6, 1973	Retrieved March 31, 2003, from http://www.mfa.gov.il
Meir 5	Statement to the Knesset by Prime Minister Meir, October 16, 1973	Retrieved March 31, 2003, from http://www.mfa.gov.il

Appendix I (Continued)

Identifier	Vision Statement	Source
Meir 6	Statement to the Knesset by Prime Minister Meir, October 23, 1973	Retrieved March 31, 2003, from http://www.mfa.gov.il
Thatcher 1	House of Commons speech: Economic censure, July 27, 1981	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Thatcher 2	House of Commons speech, Debate on address, November 4, 1981	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Thatcher 3	House of Commons speech: Economic censure, November 28, 1981	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Thatcher 4	House of Commons speech: Falkland Islands, April 3, 1982	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Thatcher 5	House of Commons speech: Falkland Islands, April 14, 1982	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Thatcher 6	House of Commons speech: Falkland Islands, May 20, 1982	Retrieved November 27, 2002, from http://www.margaretthatcher.org
Bush 1	Radio address of the president to the nation, August 18, 2001	Retrieved November 26, 2002, from http://www.whitehouse.gov/news/releases/2001/08
Bush 2	President urges quick passage of education plan in radio address, September 1, 2001	Retrieved November 26, 2002, from http://www.whitehouse.gov/news/releases/2001/09

Appendix I (Continued)

Identifier	Vision Statement	Source
Bush 3	President emphasizes education reform in radio address, September 8, 2001	Retrieved November 26, 2002, from http://www.whitehouse.gov/news/releases/2001/09
Bush 4	Statement by the president in his address to the nation, September 11, 2001	Retrieved October 27, 2002, from http://www.whitehouse.gov/news/releases/2001/09
Bush 5	Address to a joint session of Congress and the American people, September 20, 2001	Retrieved October 27, 2002, from http://www.whitehouse.gov/news/releases/2001/09
Bush 6	Radio address: Despite challenges, economy fundamentally strong, September 22, 2001	Retrieved November 26, 2002, from http://www.whitehouse.gov/news/releases/2001/09
