

STATE, TRAIT, OR DESIGN? A CRITICAL EXAMINATION OF
ASSUMPTIONS UNDERLYING REMOTE ASSESSMENT

By

SARKIS ARMEN MAHDASIAN

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

WASHINGTON STATE UNIVERSITY
Department of Political Science

DECEMBER 2002

© Copyright by SARKIS ARMEN MAHDASIAN, 2002
All Rights Reserved

© Copyright by SARKIS ARMEN MAHDASIAN, 2002
All Rights Reserved

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of SARKIS ARMEN MAHDASIAN find it satisfactory and recommend that it be accepted.

Matthew Cottam

(Chair)

David Nice

Thomas Preston

Michael D Young

ACKNOWLEDGMENTS

I must first thank Margaret Hermann, Robert Woyach, and Michael Young of Social Science Automation, Inc., whose generous support in both data and company resources made this dissertation possible. A special note of appreciation goes to Michael Young, who not only made these resources available, but also served as a member of this dissertation committee. His direction and insight have proven extremely valuable in transforming a disparate collection of hunches and skepticisms into this finished product. Margaret Hermann's theory of leadership analysis serves as the basis for this study and on many opportunities has furthered my understanding of the theory behind the technique and its application. That being said, any error in the interpretation of the technique or the resulting data found within these pages is my responsibility alone.

I have had the good fortune of working closely with many of the faculty at Washington State's Department of Political Science. Most influential has been Martha Cottam, the chair of this dissertation committee. I could not have had a better mentor and am truly thankful for all that I have learned from her. Martha's patience over the years it has taken to complete this project has seen no end and her advice and encouragement at various points along the way were instrumental in its completion. David Nice's command of the literature across the subfields of political science has helped a student of International Relations begin to open the door to American politics and his advice on this dissertation is greatly appreciated. I have Tom Preston to thank for my introduction to Leadership Trait Analysis. He taught me just enough about this topic to allow me to ask the questions found in this study but, more importantly, allowed me the freedom to arrive at my own answers. Outside of my committee I have been fortunate to have a

number of excellent faculty mentors including Bob Jackson and Otto Marenin, both of whom have helped me become a better teacher. I would also like to thank Nicholas Lovrich for his unlimited support keeping both myself and my wife employed with research opportunities during the Summer months in Pullman. Finally, I would like to thank Kunihiko Imai and Peggy James, my undergraduate professors at the University of Wisconsin-Parkside, two of the most dedicated teachers I have ever met in my many years of college.

The support staff of the Political Science department has been my saving grace throughout my entire graduate career at Washington State. Cynthia Avery, Lisa Janowski and Ruth Self all deserve medals of honor for helping me navigate the bureaucratic nightmare located at French Ad. Of special mention is Diane Berger, who undoubtedly spent too much of her own time helping me get copies of my chapters to committee members, filing the multitude of forms required to complete this dissertation and keeping me informed of deadlines. There is no way I could have completed this dissertation without her help.

It would be impossible to survive graduate school without many good friends and fortunately for me they were never in short supply. John Neiswender has taught me more about the important things in life in six years than I would have expected to learn in a lifetime, and I suspect that this is only the beginning. Bill Schreckhise has been a great friend and gracious winner whether it be on the racquetball court, the pool table, or pitching pennies on the seventh floor of Johnson Tower. You can't make it through this process without a lot of humor and I have Chris Hanks to thank for keeping me laughing every day since I first met him. Michael Gaffney has been a good friend and a great role model, I hope to someday be as good a father, husband and friend as Mike. I would also like to thank "Big" Vahan, Ted Jones, Dave Murphy, Brian Maxwell, Gray Worhtma, Frank Bilotta, Paul Painter and Dave Sedlick.

I would never have made it to where I am today without the unconditional support of my family. They have never lost faith in my ability to complete this undertaking, even when I thought it was time to move on. I see in my parents my best strengths and will be fortunate to do half as well in my own life. I must also thank my older brother, Vahan. He spent twenty years making my life miserable only to turn around and be the best friend I could ever ask for.

Above all I would like to thank my wife Kathleen. How anyone could put up with my insanity the last few years is completely beyond my comprehension. Her ability to get me to laugh when I was buried under reams of data helped keep me sane throughout this process. She has inspired me to accomplish more than I ever thought I could accomplish on my own.

STATE, TRAIT, OR DESIGN? A CRITICAL EXAMINATION OF
ASSUMPTIONS UNDERLYING REMOTE ASSESSMENT

Abstract

By Sarkis Armen Mahdasian, Ph.D.
Washington State University
December 2002

Chair: Martha L. Cottam

Scholars of international relations interested in better understanding the role individuals play in the formulation of their nation's foreign policy are hampered by the lack of accessibility of world leaders. Because such leader's have little incentive to make themselves available for a battery of psychological tests, political psychologists have developed remote assessment techniques to overcome this obstacle. In many cases these techniques, developed over the past 30 years, have provided much-needed rigor to the study of foreign policy decision-making. Even with this added rigor, such techniques face criticism on grounds of inter-coder reliability, validity, and access to sufficient material upon which leader profiles are generated. With the development in recent years of automated content analysis software, many of these criticisms have become irrelevant by eliminating reliability issues and through dramatic reductions in the costs associated with coding voluminous amounts of data.

However, the desire to charge forward and code all available data must be tempered by the need to revisit the assumptions underlying these techniques. This dissertation conducts a critical review of one of these techniques, Leadership Trait Analysis, developed by Margaret Hermann. Using the Profiler Plus automated content analysis software developed by Social

Science Automation, all public utterances of four U.S. President's first term in office are coded for each of the seven leadership traits of the Leadership Trait Analysis technique.

Four questions are asked with significant consequences for future research using both the Leadership Trait Analysis technique as well as other remote assessment techniques. The findings confirm the approach taken by Hermann that spontaneous utterances differ significantly from prepared statements. The data also show that the "person-in-context" approach taken in recent years best describes the expressed personality of the four leaders with trait scores varying significantly across statement topic and across time. Finally, the variation across time and topic place minimum sample sizes of previous research in question. This dissertation concludes with suggestions for future research necessary before embarking on a full-scale analysis of world leaders and their role in international affairs.

TABLE OF CONTENTS

	Page
ABSTRACT	vi
LIST OF TABLES	ix
LIST OF FIGURES	xi
CHAPTER	
1. INTRODUCTION	1
2. RESEARCH DESIGN	47
3. DESIGN OR REALITY? DIFFERENCES IN TRAIT SCORES BY SPONTANEITY	55
4. STATE VERSUS TRAIT I: TRAIT SCORE DIFFERENCE BY RESPONSE TOPIC	87
5. STATE VERSUS TRAIT II: TRAIT SCORE STABILITY ACROSS TIME	104
6. SAMPLE SIZE REQUIREMENTS FOR RELIABLE TRAIT SCORE CREATION	121
7. CONCLUSION	127
BIBLIOGRAPHY	159
APPENDIX	
A. DATA TABLES	166
B. FIGURES	209

LIST OF TABLES

1. Spontaneity ANOVA for George Bush, Sr.	64
2. Population Mean Trait Scores and Comparison Ranges	66
3. Substantive Difference in Trait Scores by Spontaneity Type for George Bush, Sr.	67
4. Spontaneity ANOVA for Bill Clinton	68
5. Substantive Difference in Trait Scores by Spontaneity Type for Bill Clinton	69
6. Spontaneity ANOVA for Lyndon Johnson	70
7. Substantive Difference in Trait Scores by Spontaneity Type for Lyndon Johnson.	71
8. Spontaneity ANOVA for Harry Truman	72
9. Substantive Difference in Trait Scores by Spontaneity Type for Harry Truman	73
10. Summary of Statistically Significant Variation on Spontaneity	74
11. Direction and Magnitude of Change between Trait Scores for Prepared and Spontaneous Material	74
12. Substantive Shifts for Belief in One's Own Ability to Control Events	77
13. Substantive Shifts for Conceptual Complexity	78
14. Substantive Shifts for Distrust of Others	78
15. Substantive Shifts for In-group Bias	78
16. Substantive Shifts for Need for Power	79
17. Substantive Shifts for Self Confidence	79
18. Substantive Shifts for Task Orientation	80
19. Comparison of Trait Classifications based on Different Population Material	81

20. Correlations between Spontaneous and Prepared data by Year and Quarter	82
21. Results of ANOVA by President on all Spontaneous Data	84
22. Response Topic ANOVA for George Bush, Sr.	90
23. Response Topic ANOVA for Bill Clinton	91
24. Response Topic ANOVA for Lyndon Johnson	93
25. Response Topic ANOVA for Harry Truman	94
26. Summary of Statistically Significant Variation on Response Topic	96
27. Population Mean Trait Scores and Comparison Ranges	92
28. Substantive Variation by Response Topic for George Bush, Sr.	98
29. Substantive Variation by Response Topic for Bill Clinton	98
30. Substantive Variation by Response Topic for Lyndon Johnson	99
31. Substantive Variation by Response Topic for Harry Truman	99
32. ANOVA Probability findings for variation in trait scores by President by each topic	102
33. Ordered mean scores for President Bush's traits using 50 response samples from Conflict responses	122
34. Change in classifications for 50 response samples drawn from all Bush responses where topic is conflict	123
35. Mean scores for President Bush's traits using 50 response samples (of at least 100 words)	124
36. Change in classifications for 50 response samples drawn from all Bush responses where topic is conflict and word count is greater than 99	125
37. Conceptual Complexity and Self-Confidence scores	143
38. Response Frequency by day for George Bush, Sr.	149
39. Coding Opportunities for World Leaders in Comparison Population	155

LIST OF FIGURES

1. Trait Score Differences by Spontaneity for Bush	65
2. Trait Score Differences by Spontaneity for Clinton	68
3. Trait Score Differences by Spontaneity for Johnson	70
4. Trait Score Differences by Spontaneity for Truman	72
5. Bush's Belief in One's Own Ability to Control Events by Year	107
6. Bush's Conceptual Complexity by Year	108
7. Task Orientation by Year for All Presidents	110
8. Need for Power by Year for All Presidents	110
9. Self Confidence by Year for All Presidents	111
10. Bush's Belief in One's Own Ability to Control Events by Quarter	112
11. Bush's Conceptual Complexity on Economic Responses by Quarter	113
12. Bush's Task Orientation on Economic Responses by Quarter	113
13. Bush's Belief in One's Own Ability to Control Events for Conflict Responses by Day	116
14. Clinton's Belief in One's Own Ability to Control Events for Conflict Responses by Day	116
15. Johnson's Belief in One's Own Ability to Control Events for Conflict Responses by Day	117
16. Truman's Belief in One's Own Ability to Control Events for Conflict Responses by Day	117
17. Conceptual Complexity for All Presidents on Conflict Responses aggagated by Year	139
18. Need for Power for All Presidents on Economic Responses aggagated by Year	140

19. Need for Power for All Presidents on Economic Responses aggegated by Quarter	140
20. Conceptual Complexity Scores for Truman Aggregated by Day	151
21. Conceptual Complexity Scores for Truman, Aggregated by Day (excluding days with less than 10 coding opportunities)	152
22. Conceptual Complexity Scores for Truman, Aggregated by Day (excluding days with less than 25 total opportunities)	152
23. Conceptual Complexity Scores for Truman Aggregated by Quarter	153

Dedication

To mom, dad and Vahan.

Your support will never be forgotten.

Chapter 1

Introduction

Adolf Hitler, Winston Churchill, Pol Pot, Franklin Delano Roosevelt, Margaret Thatcher. Each of these leaders played a critical role in world history over the years. Some are looked at from history's perspective as great, while others are seen as evil despots. What if these people had never made it to their positions of power? What if Hitler's Nazi party had been unable to take control of the German Reichstag in 1933? Had he not been in control of the German government, countless lives may have been spared and the devastation of Europe might have never happened. The Holocaust, the division of Europe into East and West, all may have been avoided had this one man never taken office. While Hitler is the 20th century's classic example of a single individual causing great harm to the world, fortunately we have had leaders who have brought great good to the world as well. What would American race relations be like had Martin Luther King Jr. never been born? Would we still have separate drinking fountains for blacks and whites? What if Abraham Lincoln never decided to enter political life, would George McClellan, Lincoln's opponent in the 1864 presidential election, have abolished slavery? If not McClellan, how much time would it have taken before the practice of slavery was ended in this country? And had slavery not been abolished, would we still have had the civil war? Questions such as these are endless if we think of the consequences of replacing one individual for another throughout history. Yet when studying international politics, the dominant approach argues that individuals simply do not matter.

While the majority of people view world events as the result of the actions of world leaders, there are those who argue that there is no such thing as a great man, only situations

which lead such people to greatness. Constellations of factors completely irrelevant to the leader in office are what actually produce the “great leader”, they just happen to be in the right place at the right time. Abraham Lincoln would have been a mediocre President had he not been in office during the civil war, (Skowronek, 1995). Adolf Hitler was only elected because the people of Germany needed a strong leader to move them beyond the misery brought on by the Treaty of Versailles. The Reverend Martin Luther King Jr., would simply be another preacher had the civil rights movement not reached a point where it was ready to take on a national form. Rather than grand figures moving politics, it is politics that is responsible for creating the grand figures.

The most dominant theory of international relations, the neo-realist approach, argues that individuals are of no consequence to outcomes in international affairs. For the neo-realists, all nations of the world follow a single rule - maximize your power. No matter what the personal predilections of any government leader might be, if they fail to follow this rule their country will not survive for very long (Morgenthau, 1948). Like survival of the fittest, international politics revolves around strength, namely the ability to protect one’s self and get what you want from others. According to this perspective, if it had not been Adolf Hitler who drove Europe to war in the 1940s, it would have been whoever else had been elected. Power maximization was behind the outbreak of World War II - the leader behind the helm was inconsequential. Germany’s power had reached a point where if it had not expanded to overtake others, an alliance of competing world powers would have formed to eliminate its potential to do so.

A competing explanation for the study of international politics argues that individuals can not be ignored as the neo-realists claim. Falling under the general description of foreign policy decision-making, many of these approaches utilize psychological constructs to explain why events occur in international relations. For such scholars, studies which rely solely on the neo-

realist, power politics perspective are insufficient and the detailed analysis of individual leaders is necessary to answer questions that the neo-realists are unable to address. Most importantly of which, the neo-realists can only explain past events and not predict future ones. While explanation is an important component of understanding international affairs, many scholars (and all practitioners) would argue that without the ability to predict future outcomes, approaches to politics are of little value. The differences between these two perspectives can be illustrated through the examination of a critical event in international politics: the 1990, Iraqi invasion of Kuwait. This example will shed light on both the differences between the two approaches and the reason why the neo-realist perspective is an insufficient tool for the analysis of foreign policy.

As mentioned above, neo-realism argues that the most important factor in the explanation of international behavior is that all states seek to maximize their power. Whether or not war will break out between two or more countries is simply a product of the relative power levels held between those states. If country A is strong enough to attack country B and country A will maximize its power by doing so, then country B had better be prepared for an invasion. In the case of Iraq and Kuwait, Iraq was disproportionately more powerful than Kuwait and saw an invasion as a way to dramatically increase their oil producing capacity as well as provide an access port to the Persian Gulf. The increase in oil production would bring Iraq much needed foreign reserves that it could use to increase its military might - thus maximizing its power. In addition to revenue gains from oil production, access to the Persian Gulf would allow Iraq to increase both its economic capacity as well as its military capacity through an additional avenue for the shipment of both troops and goods. This “power politics” approach to international relations has been around for thousands of years¹ and is the dominant perspective because it *does*

¹ The Realist approach and its more recent derivatives such as Neorealism and Structural Realism can be traced back at least to Thucydides and his History of the Peloponnesian War.

explain much of what happens in the realm of international politics, and does so without much complication. After all, the one rule - all states seek to maximize their power – explains everything. A brief analysis of speeches by political leaders throughout the world will surely confirm that leaders talk about power frequently and most use some sort of power politics mindset when thinking of how to deal with other states in the international system.

For those looking for a simple and parsimonious theory which explains much of what happened in past events, neo-realism is an impressive candidate. Unfortunately, simple models such as the neo-realist approach sometimes confuse issues in their drive for parsimony. The neo-realist approach is not only limited to post-hoc analysis of events but is also incapable of producing predictions of future behavior. After all, while the above analysis shows that Iraq invaded Kuwait because it was more powerful, it does not tell us why Iraq chose the summer of 1990 to do so. Iraq had the capacity to invade Kuwait prior to that date, why did it choose that point in time for an invasion? Why was Iraq the only nation in the Middle East who attempted an invasion of Kuwait? Saudi Arabia was clearly capable of overpowering the much weaker Kuwaiti armed forces yet it took no hostile action. Also, why did Iraq choose Kuwait? Did an invasion of Kuwait lead to the greatest increase in national power with the least amount of power expended in the process? If so, how is that determined? Neo-realists rely on some objective version of the cost benefit analysis process that assumes that Iraq (note, using the neo-realist approach necessitates viewing Iraq as a single entity, without reference to Saddam Hussein) views the world and its conception of power just as the United States does. Even if the issues listed above could be addressed by the neo-realists, how can a power politics approach answer why Iraq continued to remain in Kuwait after the United States threatened it with war? Surely the United States is a more powerful nation than Iraq and Iraq was bound to suffer defeat, even if it

was able to impose significant costs on the United States in the process. Surely this can not be an effective way to increase one's power. Limited to state-level, power oriented explanations, the neo-realist approach is incapable of addressing these issues. Finding answers to such questions require scholars to move beyond the neo-realist approach.

The answer to these questions can be found by an examination of the decision-making approach to international relations. Asking why Iraq invaded Kuwait in the summer of 1990 from the foreign policy decision-making perspective, a number of answers immediately come to mind. Saddam Hussein had clear intentions to increase his stature and role as a regional power in the Middle East. While this sounds similar to the power politics rule of the neo-realists, it is more precise by identifying differences between specific leaders. Saudi Arabia may have had the capacity to invade Kuwait, but it possessed an entirely different leadership who did not possess the same intentions as that of the Iraqi leader, hence the Saudis did not invade Kuwait. Perceptions also come into play with the foreign policy decision-making school. With Iraqi troops amassing at the Kuwaiti border in June of 1990, Saddam Hussein met with April Glaspie, the U.S. Ambassador to Iraq. According to an interview between Hussein and Glaspie released by Iraq to the New York Times, Glaspie gave Hussein the impression that the United States would look the other way if Iraq chose to invade Kuwait (New York Times, 1990). Later, when the United States threatened military action against Iraq if it did not withdraw from Kuwait, Saddam Hussein perceived the United States to be a paper tiger, unwilling to suffer real casualties as occurred in the war in Vietnam. Hussein miscalculated the willingness of the American public to be mobilized behind their political leaders (especially in the early stages of a conflict before the media can bombard the public with images of the reality of war). More importantly, he miscalculated the ability of the United States to assemble a multinational

coalition against Iraq, providing both international legitimacy and additional resources to effectively decimate its armed forces. Each of these explanations rely on an understanding of certain psychological characteristics of specific individuals within governments as well as knowledge regarding the specific actions taken by these people. Without examining the individuals whose actions make the history we are interested in studying, significant questions are left unanswered (or, even worse, answered incorrectly) by power politics approaches to international affairs.

Because individual differences do make a difference and are essential components of understanding both how political dramas unfold and predicting future outcomes, the benefits received from a more complete understanding of the ways individuals interact with political processes are of considerable value. Could the Gulf War have been prevented had the United States sent clearer signals to Saddam Hussein regarding its position on Kuwait? Was there any way of identifying Hitler *before* he committed such horrible acts and brought war to all of Europe?

While the focus of such counterfactuals often centers on the tragedies and horrors of world history, lessons may also be learned from positive developments throughout history. For example, which characteristics of those we call “great” leaders make them so effective in their roles? If we can identify what makes a leader great, can we use that to better choose our political representatives? Critics of presidential campaigns abound, many calling for serious reform of the electoral system. Would the use of leadership assessment help us choose better leaders? In cases of other countries where the selection of leaders is outside our control, can we better deal with those that are elected to prevent conflict and increase cooperation between states? The oft quoted

lesson learned from World War II is not to appease leaders who are bent on further domination². If this lesson is truly accurate, knowing when a leader is bent on world domination becomes critical in knowing how to best deal with other countries. Ignoring the characteristics that differentiate one leader from another will ensure that our understanding of political processes is incomplete.

Foreign Policy Decision Making

While the discussion to this point has presented the foreign policy decision-making approach as a single entity, this is somewhat misleading. Different techniques range from the study of dreams to a leader's childhood and family life to the word choices they make in their speeches and interviews³. While each technique is organized under the common label of foreign policy decision-making (or simply called political psychology), they are far from a single unified method of analysis. Before moving on to the specific method used in this study, a brief overview of the approach will be useful.

In 1954, three scholars produced a seminal work in political psychology. While not the first research using psychology to explain political behavior, Richard Snyder, H.W. Bruck and Burton Sapin provide us with a useful starting point for our coverage of the foreign policy decision-making approach to international politics. In the authors' words, "[our] framework borrowed and combined insights and conceptual guides from sociology and social psychology that had not hitherto been applied systematically to foreign policy analysis" (1963, p. 1). Snyder, Bruck and Sapin identified the importance of examining the leader's "definition of the situation".

² See Khong (1994) *Analogies at War* for an excellent examination of the role played by the lessons of history in future policy decisions.

³ For a comprehensive and near-exhaustive examination of all aspects of foreign policy decision-making see Vertzberger (1990) *The World in Their Minds: Information Processing, Cognition, and Perception in Foreign Policy Decisionmaking*

As noted above in the criticism of the neo-realist perspective above, without knowing Saddam Hussein's perception of the credibility of threats by the United States, we are left to assume perceptions based on an "objective" interpretation of the events at hand. Yet, the actual decisions made and actions taken by such leaders are not based on a scholar's interpretation of the credibility of a threat, they are based on that individual leader's interpretation. Snyder and his colleagues called for investigating the perceptions held by those who control the foreign policy apparatus. Their work is also significant in that their understanding of the definition of the situation extends further than many of the studies conducted under the foreign policy decision-making school today. Often, critics of the foreign policy decision-making approach argue that such studies are reductionist because they ignore the surrounding environment, looking solely at an individual leader as if they existed in a vacuum. If the study of individual influence in the foreign policy process is needed because systemic approaches ignore critical information, then studies that focus solely on the leader are also flawed due to the absence of critical systemic factors necessary to explain why the environment in which leaders find themselves changes. Snyder, Bruck and Sapin's definition of the situation sought to bridge the gap between the two approaches by looking at not only leaders' beliefs and experiences, but also the environment in which the leader exists.

"It is difficult to see how we can account for specific actions and continuities of policies without trying to discover how their operating environment is perceived by those responsible for their choices, how particular situations are structured, what values and norms are applied to certain kinds of problems, what matters are selected for attention, and how their past experience conditions present responses." (p. 5)

Snyder, Bruck and Sapin set the stage for decades of research which has evolved into a fairly coherent and organized research program. Brian Ripley (1993) identified the core

components of this foreign policy decision-making approach in an excellent review, in it he identified the defining tenets of the approach. The following are a sample of those tenets that are relevant to this study.

- Foreign policy elites acting on behalf of states and non-state institutions are the primary actors in world politics.
- Foreign policy elites act on the basis of their definition of the situation.
- Information is the primary currency of international relations.
- Policy prescriptions involve efforts to compensate for individual misperception and organizational pathologies. (Ripley 1993, p. 406).

The first point is the critical distinction demonstrated earlier with the comparison of how neo-realists and foreign policy decision-making analysts view the international system. While neo-realists are intent on studying the state and its power, those who integrate psychology into their study of foreign policy are focused on the individuals who represent that state. While it is, at times, useful to relegate government to the abstract entity we call the state, proponents of the foreign policy decision-making school argue that because the state is made up of individuals, it is individuals who must necessarily be part of any analysis of state action. Ripley quotes Mandelbaum by referring to this perspective as “an ‘inside out’ explanation of state behavior (based on internal features of the state) rather than an ‘outside in’ explanation (based on the constraints and opportunities imposed on states by the global system.” (Mandelbaum 1988 in Ripley) The second tenet listed above is taken directly from Snyder, Bruck, and Sapin and reiterates the importance their analysis holds in the foreign policy decision-making literature. That information is the primary currency of international relations shows the extent to which advocates of the foreign policy decision-making approach differ from the neo-realists. Power is everything in the neo-realist perspective, while the way information is “gathered, processed, and

remembered by individuals...” (Ripley, p. 410) is critical to the decision-making approach. The neo-realist assumptions of perfect information and rational interpretations of that information are thrown out the window when individual leaders become the focus of analysis. One need not look to international relations to see how far from reality such an assumption actually is. How many times have you received a task at work where you interpreted it completely differently than the person who assigned you the task? Whether it be because of poor communication (your boss gave you a one sentence description of the task), different perspectives (you used to work at a company where reports were always produced in a certain format), or the different roles you and your supervisor have within the organization (your view of how a task should be accomplished may differ based on your priorities) - different perceptions lead to different outcomes. The earlier discussion of the Gulf War shows how Saddam Hussein and George Bush were able to look at the same situation and produce two completely different projections of the outcome. Hussein thought the United States would suffer serious casualties in the desert environment and be forced to withdraw. Bush, on the other hand, was aware of the advantage provided by the aerial superiority of the United States’ military and was able to follow through on the threat to retaliate against Iraq’s invasion. While the government and the media may declare Hussein to be an insane, 1990’s version of Adolf Hitler, his refusal to withdraw from Kuwait can be explained in large part by access to different information and processing that information in a way foreign to that of the way most people in the United States think.

The final tenet taken from Ripley is that the foreign policy decision-making approach believes that policy can be improved through careful analysis of the individuals responsible for the foreign policy actions of the world’s nations. Robert Jervis’ classic study, *Perception and Misperception in International Relations* (1976), showed how simple differences in the ways

competing groups perceived their environment led to costly struggles where many lives were lost. Rather than view human beings as inconsequential as the neo-realists do, decision-making researchers see political leaders as critical components of the foreign policy process and capable of making changes to better serve their nation's interests.

“Human decision-makers are not locked into a preordained march of history or constrained by a dominant global system. Instead, decision-makers select projects in accordance with their diagnosis of policy problems, some attention to goals, and with an eye to available resources. Sometimes the decision process fails. While there may be no ‘ideal’ decision process, some conditions foster relatively ineffective outcomes because of biased problem diagnosis, systematic information distortion, and a failure to anticipate the consequences of a policy.” (Ripley, p. 411).

It is exactly this area where the foreign policy decision-making school can offer an alternative to the closed explanations of the neo-realists and provide additional understanding that is beyond the limitations of such an approach. While Snyder, Bruck and Sapin identified the importance of studying the role individual policy makers play in the formulation of foreign policy, this still leaves us with a critical dilemma: how do we get access to the leaders we wish to study? Classified documents, which some may argue is essential to understanding the “true” thoughts and beliefs of politicians, are for the most part out of the public realm for decades. Even more difficult is access to the leaders themselves. Not many leaders would be willing to sit down with psychiatrists for a battery of tests used to assess their psychological attributes.

Numerous methods have been developed since the pioneering work by Snyder, Bruck and Sapin in an attempt to solve this very problem. Michael Young and Mark Schafer (1998), review four of these methods that have used at-a-distance assessment techniques to overcome the lack of access to world leaders. These methods are (1) the Operational Code technique, (2) Image theory, (3) Cognitive mapping, and (4) Leadership Trait Analysis.

The first method of analysis used to understand and predict foreign policy behavior reviewed by Young and Schafer is Operational Code analysis. Nathan Leites conducted the first Operational Code studies in 1953, when he used the technique to identify the belief systems of the members of the Soviet Politburo. Later studies built from Leites work by refining the philosophical and instrumental components of a leader's Operational Code into a total of ten indices used for comparison (George 1969, 1979). The philosophical components deal with a leader's perception of the international environment. The leader's scores for each of these components was originally based on the assumed answers the leader would give based on the following five questions:

1. Is the political universe essentially one of harmony or conflict?
2. What is the fundamental character of one's political opponents?
3. What are the prospects for the eventual realization of one's fundamental political values and aspirations?
4. Is the political future predictable? In what sense and to what extent?
5. How much control or mastery can one have over historical development? (Young and Schafer 1998, p. 70)

The instrumental components focus on a leader's strategy for dealing with the international environment. The instrumental components of the operational code were derived by answers to the following questions:

1. What is the best approach for selecting goals or objectives for political action?
2. How are the goals of action pursued most effectively?
3. How are the risks of political action calculated, controlled and accepted?
4. What is the best timing for action to advance one's interests?
5. What is the utility and role of different means or tactics for advancing one's interests? (Young and Schafer 1998, p. 70)

Taken together, these ten indices are used to produce an operational code for the leader that reflects "the leader's worldview and response repertoire produced by socialization experiences and personality traits" (Walker and Schafer 2001, p. 530).

While early attempts at producing an operational code for leaders was done by studying as much about a leader as was available (through their actions, comments made of them by others, and their statements and memoirs) this proved to be more subjective than was desired. One analyst's conception of how Brezhnev would view the role of chance in international affairs could differ greatly than another analyst's perspective. Ole Holsti and, later, Steven Walker and Lawrence Falkowski made attempts at increasing the rigor behind operational code construction by creating a questionnaire that would be given to analysts, area experts, and policymakers familiar with a specific leader to reduce the subjective nature of the code creation by as much as possible (Holsti 1977, Walker and Falkowski 1984). Later, Walker's collaboration with Mark Schafer and Michael Young led to the elimination of third party interpretations of the leader and moved directly to an analysis of the leader's own statements. Using the Verbs in Context System (VICS), every verb based attribution made by the leader is coded for a number of characteristics and is then used to create seventeen indices used to calculate the leader's operational code (Walker, Schafer and Young, 1998, 1999).

Applications of the current method for operational code analysis have been used to both assess previous events in world history as well as predict future outcomes during crisis situations (Schafer and Crichlow 2000; Schafer and Walker 2001; Walker and Schafer 2000; Walker, Schafer and Young 1998, 1999). Knowing whether the head of a government sees the world as inherently conflictual is likely to influence whether they interpret attempts at negotiation as sincere or merely a ruse for distraction. In a crisis situation it would be critical to know what role the leader believes chance plays in international affairs (one might hope the decision maker responsible for initiating a nuclear attack does not view chance as playing the predominant role in international outcomes!). Some of the more innovative studies using the operational code

construct that are being conducted today include complex bargaining analyses between two leaders using their operational codes as a basis for strategy prediction (Walker and Schafer 2001).

The second approach reviewed by Young and Schafer is Image theory. Based on the cognitive concept of schemas, Image theory was first developed by Richard Cottam (1974) and examined the categories used by world leaders to classify the multitude of states in the international system. As a general rule, human beings are incapable of dealing with every bit of information they are exposed to on a daily basis. The only way they can function in such a complicated environment is through simplification. While there are currently over 180 states in the world, leaders do not have in-depth knowledge of each one of those 180 states. In order to think about international politics and make decisions, leaders are forced to simplify the world and thus develop categories of states. Research conducted by Martha Cottam (1986, 1992, 1994) has shown that decision makers use approximately seven categories to classify the various types of states in the international system. These include the enemy, the ally, the barbarian, the rogue, the colonial (split into two groups - colonial tied to the enemy, and colonial with no affiliation), and the degenerate. Each category has certain characteristics that allow easy placement of states in one of the seven types.

While a leader may not know much about Zimbabwe, they know enough superficial information to say it is a third world country in Africa, it is poor and not highly developed, and so on. If the leader who observes Zimbabwe comes from the United States, the leader will most likely view Zimbabwe through the colonial image (colonial states are weaker than the perceiver's state, have benign intentions towards the perceiver's state, and are perceived to be culturally inferior than the perceiver's state). Not only does the image allow the leader to place

the country in a category allowing them to answer questions about that nation, the image also acts as a filter, playing an intermediary role in the information processing process. Actions taken by Zimbabwe in the military realm are going to be perceived differently than actions taken by a state that the leader views through the rogue image. For example, if Zimbabwe decides to increase its capacity to produce fertilizers or other biologically oriented products, an official in the State Department of the United States is likely to perceive this as an attempt to increase crop yields (the colonial is seen as weak and underdeveloped and also has benign intentions towards the U.S.). If, on the other hand, Iraq decides to increase its fertilizer production capacity, that same State Department official is likely to view it as an attempt to increase its biological weapons capacity (the rogue, while weak and underdeveloped, has harmful intentions towards the official's state).

While “[o]perational code analysis grew largely out of a desire to understand and predict the behavior of an adversary, ... cognitive mapping grew out of Robert Axelrod’s desire to improve the quality of decision-making processes of the home team” (Young and Schafer, p. 74-5). One of the fundamental tenets of the cognitive approach to foreign policy analysis is that human beings are limited information processors. Many times we make decisions that are based on faulty logic or are not internally consistent with our other beliefs. Because we are overwhelmed with massive amounts of incoming information and do not always have time to process our current behavior in relation to our prior beliefs, we make decisions which may, in retrospect, seem incorrect. Axelrod (1972, 1976) intended cognitive mapping to be used to overcome these problems by explicitly charting the beliefs held by individuals or groups and their connections. Because actions taken for one reason may have unintended consequences (for example, increasing one’s defenses in an attempt to be more secure may spark a pre-emptive

strike by the adversary you were attempting to deter) Axelrod hoped to make clear the connections between a leader's intentions and the consequences of taking such action.

Constructing a cognitive map for an individual consists of converting their verb based statements into a series of nodes and connections. For example, the following statement: Chinese military exercises in the Taiwan Straits will lead to less stability in the region, would be converted into the following.

Node 1	Connection	Node 2
Chinese Military Exercises	→ (-)	Regional Stability

Which simply shows that Chinese military exercises have a negative effect on regional stability. After performing this process on all statements in a series of speeches or interviews from a leader, a complete cognitive map can be developed which can be used to understand how a leader perceives the world. Michael Young's recent work with WorldView (1994, a software package which automates the coding and graphing of a leader's cognitive map) allows for more complex relationships between nodes and can be used to predict the future behavior of a leader based on their statements in past situations. For example, if a leader has repeated instances of a negative causal connection between unchecked aggression and national security, they are more likely to respond to acts of violence (such as the Iraqi invasion of Kuwait) rather than attempt to appease the aggressor in the situation. If one is interested in knowing how a leader might respond to certain initiatives, a cognitive map can be used to predict whether they would respond favorably or negatively, and thus save policymakers from potentially making an offer that would hurt their interests.

The last of the cognitive methods reviewed by Young and Schafer is Leadership Trait

Analysis. Developed by Margaret Hermann (1983), this technique employs an at-a-distance method to assess a leader on seven different personality traits and then uses their relative scores on those traits to predict how they would react in future situations. Because this study uses Leadership Trait Analysis in its analysis, a more thorough discussion will be presented on this technique. Before discussing the technique itself, we will begin with a more general discussion of what is meant by personality and how it has been used in the political science literature.

PERSONALITY

Many people use the term ‘personality’ when describing the people they deal with in their everyday lives. Co-workers, family members, friends - if asked to describe any of these people we often use personality characteristics to describe what type of people they are or why they do the things they do. Jerry is so nice, Mary is insecure, and Larry is just too logical! In fact, with the growing use of the Meyers-Briggs Typology Indicator, millions of people have been exposed to personality assessment and how it affects their work performance (Consulting Psychologists Press, inc.). The Meyers-Briggs assessment technique classifies an individual on four spectrums (extroversion - introversion, sensing - intuition, thinking - feeling, and judging - perceiving). An interviewer asks the subject a series of true / false questions regarding their personality and then scores the test to place the individual in one of sixteen categories. Each category then provides a general description of the individual and their strengths and weaknesses in career settings. Below is a set of characteristics taken from an assessment of an individual who is classified as INTJ (introvert, intuitive, thinking, judging):

- Insightful, conceptual, and creative;
- Rational, detached, and objectively critical;

- Likely to have a clear vision of future possibilities;
- Apt to enjoy complex challenges;
- Likely to value knowledge and competence; apply high standards to themselves and others;
- Independent, trust their own judgments and perceptions more than those of others;
- Usually seen by others as private, reserved, and hard to know. (Consulting Psychologists Press, inc.)

In addition to career planning by corporations and government agencies, personality has been used by political scientists as a means to assess leaders and their behavior. One of the most famous uses of personality (and the most criticized) is James David Barber's *Presidential Character*, first published in 1972. Firmly believing that who occupies the White House has a significant effect on U.S. politics, Barber's stated goal is "to help citizens and those who advise them cut through the confusion and get at some clear criteria for choosing Presidents" (1985, p.1). Barber argues that a politician's character, worldview, and political style all explain how they will deal with the stress associated with being President of the United States. The core of Barber's argument is that political figures can be classified into one of four character types which are a function of two variables: a) how active the individual is and b) whether or not he gives the impression he enjoys his political life (p. 4). Depending on the classification of the individual on each variable, Barber places Presidents into one of the following four categories: a) active-positive; b) active-negative; c) passive-positive; and d) passive-negative.

After placing the President into one of these categories, Barber identifies the political consequences of this Presidential character. Among other things, active-negative Presidents continually focus on their power relationships with others, deny themselves self-gratification, have a perennial temptation to fight or quit, views the world as a dangerous place, and

approaches his work as an obligation or duty (p. 81-83). Barber argues that electing this type of person to the office of the Presidency may very well lead to disaster. Active-negative Presidents will continue support for policies long after they have been proven to be counterproductive and place a strong emotional attachment in such policies. The most fitting example of a President who fits into this character identification is Lyndon Johnson. His handling of the war in Vietnam, Barber argues, is characteristic of the predictions listed above for active-negative Presidents. Barber's work on Presidential Character receives much attention for a number of reasons, however one of the most controversial statements in an edition released before Nixon's re-election to a second term in office was that he was likely to suffer a catastrophic fall due to his personality. In the years since Barber has released numerous editions of his text with critics continuing their attacks of his work⁴.

While the focus of Barber's study is the political elite, Stanley Renshon (1974), uses personality as a way to explain how the general public becomes involved in political life. Focusing on the concept of "personal control", he argues that within all people is a "basic need to gain control over [their] physical and psychological life-space" (p. 1). Renshon acknowledges that there are many other factors which serve as relevant explanations for political action, but that his purpose is to "trace [the] relationship between human psychological needs and political life, by focusing on one specific need, the need for personal control" (p. 3). In order to explore this relationship, Renshon interviewed 300 undergraduate students in an attempt to test a number of hypotheses regarding personal control. These hypotheses deal with the family structure in which individuals are raised, the nature of the relationship the individual has with their parents, and the consequences of higher or lower levels of personal control on both political and non-

⁴ One of the more reasoned and productive critiques of Barber's work can be found in George and George (1998), *Presidential Personality and Performance*.

political areas of the individual's life. Renshon's examination of these hypotheses leads to a number of conclusions regarding citizen participation and satisfaction in the political life of their country.

Greenstein's use of personality is most relevant to the application of personality used in this analysis because of his concern with *when* personality matters. Greenstein's model is particularly attractive to this study because it bridges the gap between competing approaches that argue that either the system or the individual is all that matters. While appreciating the role individuals play in the formulation of government policy, Greenstein focuses on situations that make the personal characteristics of the leader more or less relevant to the behavioral outputs of government. In *Personality and Politics* (1969), Greenstein presents a response to the prevailing criticisms of the use of personality in studies of political behavior.

Greenstein's research identified five of the more prevalent critiques of personality studies. The first is that personality characteristics tend to be randomly distributed within institutional roles and therefore is not important for analysis. He argues that this critique is invalid because of two points: first, the empirical evidence shows that these characteristics are not randomly distributed; and second, that even if you assumed random distribution, you still would not have random distribution throughout the various small groups within government institutions. The second critique he addresses concerns the relative importance of people's personality characteristics versus their social characteristics. Critics argue that social characteristics are more important, thus eliminating the need to study personality. Greenstein says that this point is also incorrect due to a misconceptualization of the terms. The social characteristics are actually the social environment and merely illustrate the situation in which the individuals find themselves. Rather than being more important, the social environment

simply sets the stage in which the individual responds (with the personality having an effect upon those responses).

The next three critiques Greenstein addresses are valuable in the sense that they serve as indicators of *when* personality matters. The first of these three deals with the ability of actors to shape events. Here, Greenstein points out the difference between action dispensability and actor dispensability (which is covered in the next critique). Action dispensability deals with the degree to which individual actions can shape the outcomes of events. Greenstein points out three variables which will affect the degree of action dispensability: 1) how vulnerable the environment is to restructuring; 2) the strategic location of the actor within the environment; and 3) the abilities of the actor. These items enable us to answer the question “What are the circumstances under which the actions of single individuals are likely to have a greater or lesser effect on the course of events?” (p. 42). The next critique argues that the individual personalities will have little effect on the situation because people will behave similarly when placed in similar situations (thus arguing against actor indispensability). Greenstein proposes a number of variables which will affect the degree to which personality matters, including: the level of ambiguity in the situation at hand, the demand placed on the decision-maker, and the level of affective involvement the decision-maker possesses for politics. The last class of objections identified by Greenstein deal with certain aspects of individual psychological functioning such as ego-defensive mechanisms. Critics argue that these are not important factors and, like the other critiques, point towards the inefficacy of studying personality. Greenstein responds by stating that this critique is similar to the last critique in that it requires us to ask the following question: under what circumstances do certain psychological variations have effects on political events?

While personality studies may be critiqued on methodological grounds, Greenstein excludes such criticism because, in relation to the other problems, they are the easiest to address. Methodological problems can be found throughout all the approaches to the study of politics and problems of that nature do not necessarily resign an approach to failure. Greenstein's responses to the critiques identified above address whether the study of personality and politics is a worthwhile endeavor at all, thereby providing a justification for the continued study of personality within certain bounded realms.

Greenstein continues with an analysis of three methods available for personality assessment. He argues that these methods should be seen as complementary rather than as competing because each provide the analyst with different results important to the analysis of personality and its effects on politics. The first approach is phenomenology, which "refers to the observed behavioral phenomena upon which all further analysis depends – the regular patterns of behavior that the individual or type exhibits under varying environmental conditions" (p. 65). The second approach is dynamics: where the concern shifts from phenomenological description to dynamic interpretation of the personality of the individual at hand. The last approach is the attempt at deriving the genesis of the personality. This is often the most controversial as well as the most difficult to accomplish due to the lack of data and complexities of the genesis of an individual's personality. Greenstein advocates the use of these approaches for a number of reasons: a) it organizes the task of personality diagnosis roughly in the order of degree to which an investigator's findings are likely to be able to be readily accepted by other investigators; b) order diagnostic operations in terms of their priority for the analyst of any specific pattern of political activity; c) the third and most fundamental use for this set of distinctions is theoretical classification (pp. 66-68).

Greenstein concludes with suggestions for the improvement of personality studies in the future. First, the investigator need be as explicit as possible in the formulation of hypotheses about the subject. Second, as much as is possible, the hypotheses and interpretations should be kept distinct from the observational data upon which the former are based. Third, to be imaginative regarding the creation of specific operational criteria for the acceptance or rejection of evidence to support or refute the hypotheses. Finally, personality scholars need to move towards developing standards for increasing reliability and validity claims of their analyses.

While the Myers-Briggs assessment and the authors covered in the preceding pages employ different conception of personality, there is virtually no end to the various definitions of personality found in the psychological literature. Personality, as it will be used in this study, is defined as "... a construct that is introduced to account for the regularities in an individual's behavior as he responds to diverse stimuli" (Greenstein, 1969, p.3). One of the more common methods used to describe this "construct" is to categorize individuals on various traits, stemming from the work done by Gordon Allport (1937) and Ross Stagner (1937). Traits, in essence, are abstract constructs created by observers to describe other people's actions. Without necessarily thinking about it, most people use traits to describe the personality of those they interact with. "[A]ggressive or passive, spontaneous or inhibited, may move quickly or slowly, may be impulsive or cautious, outgoing or reserved" (Forgus and Shulman, 1979, p. 11) - each of these traits are commonly used descriptions of the people we interact with in our everyday lives. Depending on the type of behavior the observer is interested in, different traits will be used to either understand or predict the behavior of another individual or group of individuals. Developed from the trait-based conception of personality, Margaret Hermann created the Leadership Trait Analysis method for personality assessment.

Leadership Trait Analysis

In 1983 Hermann produced a coding manual for the technique detailing both the theory behind each of the traits as well as the process by which researchers could conduct the measurement process. The manual identifies each of the eight personality traits Hermann has identified as being politically relevant and thus useful in the study of foreign policy behavior, including: 1) belief in one's own ability to control events, 2) conceptual complexity, 3) distrust of others, 4) in-group bias, 5) need for affiliation, 6) need for power or influence, 7) self-confidence, and 8) task orientation. Each of these traits is produced through an At-A-Distance assessment technique. Because world leaders have neither the time nor motivation to be subjected to in-depth personality assessments, we are forced to observe their personality from afar.

In developing a system to measure the personality of such individuals, Hermann made a number of choices about how to best represent a leader's personality. Hermann advocates the use of spontaneous utterances (non-prepared responses during interviews and press conferences) to prevent the ghost writing effect from biasing the results of the assessment. While some leaders spend more time than others creating the speeches they deliver at various events, most world leaders have no other option but to have a team of speechwriters producing the bulk of the words they speak at public events. By using only spontaneous utterances found in press conferences and interviews, the words spoken by the leader are presumed to be more representative of the leader's personality than those written by their advisers.

Hermann suggests a minimum of 50 responses of at least 100 words to guarantee confidence in the trait scores but recommends more if the data is available. Because there are

seven distinct traits coded with the technique, often data availability was less an issue than available labor to do the coding. Hand coding each of the seven traits for the recommended number of responses is an extremely time consuming process that falls prey to the normal criticisms of inter-coder reliability issues. Because of the massive labor requirements involved in producing personality profiles for multiple leaders, rigorous statistical studies regarding trait consistency have been limited to the conceptual complexity trait, which has the least complicated coding decisions associated with it and lent itself to automated coding techniques.

Before addressing the substantive findings of the Leadership Trait Analysis research program, a discussion of each of the seven traits is in order. The following descriptions of each trait and how they are coded from leader's spontaneous utterances are drawn from the coding manual for Leadership Trait Analysis (Hermann: 1983,1999). Material within quotes has been taken directly from the manual.

Belief in one's own ability to control events:

As its name suggests, belief in one's own ability to control events measures a leader's perception of the amount of control they have over the situations they encounter. Does a leader find that he is overwhelmed by the multitude of forces that produce outcomes in foreign affairs or does he see himself as one of those forces? Leaders who score high for this trait view themselves as an integral component in policy formation, taking a proactive role in the policy process. Those who score low simply see themselves as stewards over the process, waiting for events to occur and viewing much of their responsibility as being out of their control.

Coding for this trait is accomplished by searching the leader's responses for verbs where the subject is either the speaker themselves or a member of a group the speaker identifies with.

After locating each of these verbs, the coder then checks to see whether or not the verb indicates action taken or planned by the leader or their in-group. If an action verb is found then it is coded positive, if it is not found it is coded negative. The score for this trait is then calculated by taking the ratio of positive instances to the total number of verbs found.

Need for power or influence:

"Need for power indicates a concern for establishing, maintaining, or restoring one's power or, in other words, the desire to control, influence, or have an impact on other persons or groups". Leaders who are high in need for power will insert themselves into the political process at every opportunity. Other actors exist solely for their manipulation and are exploited for as long as they serve a useful purpose. Leaders who score lower in need for power are willing to open up the political process to allow other actors significant influence. Taking credit for accomplishments is not as important as accomplishing the goal. Allowing others to prosper is an active concern for such leaders.

When coding for need for power, all verbs where the subject is the speaker (or a group the speaker identifies with) are examined to see if any of the following conditions exist: 1) representative of strong or forceful action; 2) gives advice or assistance when not solicited; 3) attempts to regulate the behavior of another person or group; 4) tries to persuade, bribe, or argue with someone with no intention to reach agreement; 5) endeavors to impress or gain fame with an action; 6) shows concern for the subject's reputation or honor. Verbs meeting any of these conditions are coded positive and the power trait score is obtained by dividing the number of positive verbs by all verbs that have the speaker as the subject.

How Leaders Respond to Constraints

Depending on the leader's scores on these first two traits we may predict how they react to constraints in their environment. Leaders high in both belief in one's own ability to control events and need for power or influence are the most likely to challenge constraints and push the limits of what is possible. These leaders are confident of themselves and certain of both their authority as well as their knowledge of the proper course of action to take. Attempts at including other perspectives or achieving compromise are not a high priority for such leaders. Leaders who score low in both traits are more likely to respect the constraints faced in their environment and let the flow of international politics move them from one area to the next. These leaders are less likely to attempt innovative methods to overcome the problems they and their nations face in dealing with other countries. These leaders also find that building consensus and achieving compromise is an important component of governance.

Conceptual Complexity:

"Conceptual complexity is the degree of differentiation which an individual shows in describing or discussing other people, places, policies, ideas or things". Leaders low in conceptual complexity see the world as very black and white and often are driven by stereotypes of other peoples and places. Decision-making is rarely complicated for such a leader who finds little ambiguity in their environment. A leader who is high in conceptual complexity sees the many shades of gray in the people and events they encounter and often have difficulty making decisions because they are constantly searching for additional information from which to make their decision.

Coding for conceptual complexity examines all words in the leader's remarks and looks for words that exhibit high or low levels of differentiation. Examples of low differentiation

words are: always, never, absolutely, definitely. Examples of high differentiation include: sometimes, possibly, approximately, for example. The conceptual complexity score is calculated by the ratio of high complexity words to the total number of words that fit either the high or low category.

Self Confidence:

The self confidence trait measures the leader's sense of self in relation to their environment. It "... indicates one's sense of self-importance, an individual's image of his or her ability to cope adequately with objects and persons in the environment". Leader's scoring high in self-confidence pay less attention to incoming information than those who score lower on this trait. While leader's with a high degree of self confidence are more likely to make final decisions, those leader's who have low levels of this trait will tend to re-examine issues multiple times searching for the best approach. When placed in a rapidly shifting environment these leaders will often rely on those around them to provide stability.

Coding for self confidence focuses on the following pronouns: "my, myself, I, me, and mine". Whenever these pronouns are found within a leader's remarks, the coder checks for any of the following conditions: 1) represents instigation of an activity; 2) presents the self as an authority figure; or 3) is the recipient of a positive response or praise from another actor. As in the case of need for power, whenever these conditions are met the instance is coded positive. The trait score is then calculated by dividing the number of positive instances by the total number of self references in the text.

Response to Contextual Information

Combining the self confidence and conceptual complexity traits allows us to predict how

open or closed a leader will be to contextual information found in their environment. “Those whose scores on conceptual complexity are higher than their self-confidence scores are open, they are generally more pragmatic and responsive to the interests, needs, ideas, and demands of others.” These leaders are aware of their surrounding environment and act based on what they feel to be acceptable under the given conditions. These leaders are more likely to organize their decision-making structure and advisory systems in a collegial manner, allowing for dissent and to increase both the amount and variety of information presented to the group for analysis.

“Leaders whose self-confidence scores are higher than their scores on conceptual complexity tend to be closed, they are ideologues, principled and driven by causes.” Rather than being open to contrasting points of view, these leaders tend to discount any information that is not in accord with their previous thoughts and beliefs. In fact, they will often reshape incoming information in order to have it conform to their previous conceptions of the situation. While those leaders who are open to contextual information organize their advisory system in a collegial manner, leaders whose self confidence levels are higher than their conceptual complexity will structure their decision-making groups in a hierarchical manner in an attempt to control the process as much as possible. This will have the effect of limiting debate and stifling policy options that may not fit with previous conceptions of how to address the problem at hand.

An example of a leader who fits this second category is Lyndon Johnson. Studies of the decision-making process during the Vietnam War show that Johnson and his colleagues were closed off to incoming information. Any dissent from within the group was met with hostility and the adviser who challenged the group perspective found himself out of the loop. Strategies were not altered to deal with the current situation because that would have required an

acknowledgment that previous policy was ineffective. While the United States used tactics suitable for armed conflict against another large army, the reality of the situation was that they were fighting a small (and highly organized) guerilla movement who were not prone to the same weaknesses that a large army faced. The failure to shift tactics and strategy accordingly are a direct result of Johnson's failure to respond openly to incoming contextual information.

Distrust of Others:

Distrust of others examines the level of trust a leader has in other actors. Leaders high in this category often see others as competitors to their goals and aspirations. They also view the actions of others as having ulterior motives and are suspicious of others' behavior. In extreme cases, "distrust of others becomes paranoia in which there is a well-developed rationale for being suspicious of certain individuals, groups, or countries".

In coding for this trait, the coder looks for references to other actors and then checks to see whether or not the speaker is wary or distrustful of that actor. In addition, the coder checks to see if the actor is perceived as harming the interests of the speaker. To calculate the trait score for distrust of others a ratio of actors associated with distrust compared to the total number of actors referenced in the responses is calculated.

In-group Bias:

The in-group bias trait focuses on the degree to which a leader praises his own group (be it political, social, ethnic, etc...). Leaders high in in-group bias are highly nationalistic and are leery of any kind of external involvement within their. Leaders who score lower on this trait are not disinterested in protecting their group or unpatriotic, they are however, less likely to take an

"us versus them" approach when dealing with other actors.

The unit of reference when coding for this trait is any group to which the leader identifies. For each reference to such a group, the coder checks for modifiers that are favorable, suggest strength or indicate the need to maintain the group's honor and identity. The in-group bias score is calculated from the ratio between references to the in-group which possess these modifiers compared to the total number of references to the group.

Task Orientation:

The last personality trait assessed within this approach is a leader's task orientation. "Leaders have been recognized as performing two distinct functions in groups, that of moving the group toward completion of a task (solving problems) and that of maintaining group spirit and morale (building relationships)". A leader high in task focus will spend the majority of their resources on solving the problem or accomplishing the goal, irrespective of what that means for the status of the group associated with the issue. On the other hand, a leader who is low in task focus will often pay more attention to group cohesion than on ensuring the problem is tackled with optimal efficiency.

When coding for task focus, the coder searches for keywords which indicate either work on a task or other instrumental activity (coded as positive) and words which represent concern for feelings, desires, and satisfaction (coded negative). The task focus score is calculated by dividing the number of positive words found by the total number of positive and negative words.

Is the Leader Motivated by Problems or Relationships?

Depending on the scores for the three remaining traits (distrust of others, in-group bias,

and task orientation), we can identify whether a leader is motivated more by problems or whether they are motivated by relationships. “[L]eaders are driven, in general, either by an internal focus (a problem) - a particular cause, an ideology, a specific set of interests - or by the desire for a certain kind of feedback from those in their environment (a relationship) - acceptance, power, support, acclaim.” Similar to supervisors in work settings, some bosses are concerned with meeting goals or getting the immediate task accomplished and have no concern with the morale or feelings of their subordinates. Other supervisors, however, are so concerned with the feelings of their subordinates that they neglect to push them to get the task accomplished in order to maintain a friendly relationship with their colleagues. Leaders who are more focused on tasks view the world as a series of opportunities which can be used to better both their own and their group’s position in the world. Leaders who are more focused on group affiliation, on the other hand, are more focused on the groups survival and immediate well being and often view the world as a threatening place where potential threats lurk behind every corner.

Research conducted with the Leadership Trait Analysis technique

Hermann’s first analysis involving Leadership Trait Analysis and a sizable sample of leaders was done in 1980. In her analysis, Hermann moved beyond traditional applications of personality studies in foreign policy by studying more than whether aggressive leaders lead nations to war and passive leaders maintain international harmony. Hermann identified a number of personal characteristics she found to be related to foreign policy behavior in previous research. Two variables were identified which were predicted to affect the probability that the leader’s personality would be reflected in the foreign policy behavior of their respective state: a) interest in foreign affairs and b) training in foreign affairs (Hermann 1980a, p. 13).

Using the Foreign Broadcast Information Service and the New York Times, Hermann was able to collect at least fifteen verbatim interview responses for forty five heads of government from across the world. Her findings indicated that some of the variables produced high levels of significant correlation when matched with predicted behaviors for the leaders under analysis, while other variables performed less satisfactorily. Hermann concluded that these variables, including interest in foreign affairs, needed re-conceptualization of the relationship between these variables and actual foreign policy behavior (1980a, p.43).

Hermann's next study was able to examine considerably more data and studied the effect of personality characteristics on 53 separate heads of state (1984). In this study Hermann found that interest in foreign affairs was moderately inversely related to conceptual complexity, meaning that more conceptually complex leaders paid less attention to foreign affairs issues. Upon further analysis Hermann suggests that this is at least in part a product of competitive domestic political systems which require more attention be paid to domestic concerns. Hermann also examined the relationship between nationalism and the seven traits identified by Leadership Trait Analysis. Those leaders who were identified as high in nationalism were also higher than other leaders for distrust of others, self confidence, and need for power. Hermann then concluded the analysis by showing that leaders who were highly sensitive to their environment and had high interest in international affairs showed strong relationships between their personality scores and the foreign policy behavior of their states. On the other hand, leaders with little interest in foreign affairs and lower sensitivity to their environment had a poor correlation between their personal characteristics and their respective state's foreign policy behavior.

Leader-Advisory systems:

While the Leadership Trait Analysis literature discussed thus far focused on how an individual's personality affects their behavior, another thread in the research program examines how personality affects the type of advisory system a President uses to facilitate the decision-making process. Whether it be a small dictatorship or the massive governmental apparatus of the United States, leaders, no matter how powerful they are, do not do everything themselves. In the case of the U.S. presidency, the massive information processing requirements alone require a significant staff in order to make the multitude of decisions that are required of the office on a daily basis. Advisory systems play a significant role in the specific decisions that are made, especially during crisis situations. Many studies of the Cuban Missile Crisis have focused on the ExCom meetings (the secret meetings between Kennedy and his advisers dealing with the crisis) as an integral component of the successful resolution of the situation. In that case the collegial decision structure employed by Kennedy and his well timed absences from the meeting room allowed for varied discussions and prevented some of the decision-making maladies witnessed in other key foreign policy decisions⁵.

Thomas Preston (Hermann and Preston 1994, Preston 1996, 1997, 1999, 2001) has spearheaded the use of Leadership Trait Analysis in examining the types of advisory systems Presidents use to organize their office and how those structures influence the foreign policy decision-making process. "Why is it that some Presidents and advisers work together efficiently as a decision group, seek information, tolerate adviser dissent, and actively consider and debate a number of different policy options, and in other cases they do not? What is it that leads some advisory groups to consider only a narrow range of information or policy options, exclude

⁵ See Janis, Irving (1982) Groupthink: Psychological Studies of Policy Decisions and Fiascoes, for a detailed discussion of the negative consequences sometimes associated with group interactions in crisis settings.

dissenters, and limit debate to a very small inner circle around the President?" (Preston 1997, p. 192). Preston uses the leader's personality profile to assess what type of decision-making system they implement and it is this system which is used to answer the questions raised in the quote above.

Preston focuses on two of the personality traits in Leadership Trait Analysis, conceptual complexity and need for power. Combining each of these traits with a leader's level of expertise in a given policy area, Preston identified sixteen different leadership combinations that a President can embody and used that information to predict the type of advisory system structure they would use during their presidency. The various categories a leader can be placed into include the: (based on need for power) Director, Magistrate, Administrator, Delegator; and (based on conceptual complexity) Navigator, Observer, Sentinel, and the Maverick. Studying the post World War II U.S. Presidents, Preston observed Presidents Truman and Johnson to be Magistrate-Mavericks (meaning they scored high for need for power, low on conceptual complexity, and inexperienced in the foreign policy realm). He identified Presidents Eisenhower and Kennedy as Director-Navigator's (high need for power, high conceptual complexity, and experienced in the foreign policy realm); President Bush as an Administrator-Navigator (low need for power, high conceptual complexity, and experienced in the foreign policy realm while President Clinton was determined to be a Delegator-Observer (low need for power, high conceptual complexity, and inexperienced in the foreign policy realm).

Preston then conducted in-depth case studies of critical foreign policy events within each of these leaders administration's and found that his hypotheses on leadership style correlated with both the type of system each President developed, as well as the way each group handled the foreign policy crisis with which they were faced.

Spontaneity of Material used to Derive Trait Scores:

Included in a recent symposium on at-a-distance psychological assessment were two articles that looked at one of Hermann's seven traits - conceptual complexity. The first article by Brian Dille (2000) examined speeches and press conferences materials for two U.S. Presidents: Ronald Reagan and George Bush, Sr. Using a sample size of thirty documents, Dille measured authorship effects for both Operational Code indices and the conceptual complexity trait within Leadership Trait Analysis. While he found significant variation between source type (spontaneous versus prepared utterances) for the Operational Code indices, he found no significant variation for the conceptual complexity trait. Dille also looked at the topical content of the statement uttered by each President. While Bush's scores remained consistent across different topics, Reagan's complexity scores showed significant variation depending on the topic of conversation. Dille posits that this may be due to Reagan not only "performing" the words of others in his speeches but also for that of his news conferences.

Another study conducted by Dille and Michael Young (2000) examined the authorship question for two other U.S. Presidents, Jimmy Carter and Bill Clinton. In this study Dille and Young did find significant variation in conceptual complexity scores based on source type. In fact the differences between the scores in each case were of almost identical margin, leading the authors to conclude that the difference in scores may be a function of the style and setting of press conferences versus the delivery of prepared remarks. "It may be that the higher variety of issues raised in press conferences or other question-and-answer sessions naturally raises the conceptual complexity of the individual answering the multiple queries. (p. 593). The other possibility entertained by Dille and Young is that the speechwriters craft the nature of the speech

to reduce the level of complexity of the speaker in order to conform to some designated acceptable level for the public.

Conflicting results such as these may be a function of the sample size used to produce the conceptual complexity scores or it may indicate that the level of spontaneity is only relevant for some leaders. In addition, these studies have only been conducted using the conceptual complexity trait; further study of all seven traits may lead to additional insights into the nature of the authorship effect.

The preceding studies are evidence of two critical concerns for the study of foreign policy analysis. First, we can identify specific personality traits for individual leaders and they are distinguishable from one another on the basis of those traits. Second, the traits analyzed by Leadership Trait Analysis have been found to correlate with the political behavior of the states or organizations those leaders represent. In short, leaders do matter and Leadership Trait Analysis can both explain past behavior and predict future outcomes based on the hypotheses associated with the method.

Issues raised by the Leadership Trait Analysis technique

While the previous research shows the promise of Leadership Trait Analysis technique, critics have levied attacks at the method that have yet to be answered satisfactorily. Included in the 1980 issue of *International Studies Quarterly* alongside Hermann's study was a research note by Karen A. Rasler, William R. Thompson and Kathleen M. Chester (Rasler et al., 1980). Revisiting a critique of foreign policy decision-making approaches in general, Rasler and her colleagues asked whether the personality variables identified by Hermann were reflective of the leader's actual personality or whether these variables simply reflected a snapshot of the leader at

the moment the statement was given. The critique is an important one for the study of individual influence in foreign affairs because it hits at the heart of the enterprise itself: do individuals have a persistent and stable personality or does the personality traits constructed by scholars fluctuate with time and place? If stability is not found to exist, defense against neo-realist claims that individuals are insignificant becomes a much more difficult task. On the other hand, if the traits measured by Hermann's technique are stable across time and setting, then we can identify a coherent personality for each leader and match it to behavior to test its ability to explain foreign policy behavior.

Using Hermann's methodology, Rasler and her colleagues collected additional documents for ten of Hermann's 45 leaders and performed the content analysis technique on those documents. Because Rasler, Thompson and Chester were interested in trait stability over time, they split the data for each leader into two equal time periods. Their intention was to compare scores calculated from the first period to the second, assuming that "[i]f we start with the assumption that a leader's personality remains relatively constant over time, then we would expect that the interview responses for a leader in the first half of his or her administration would be correlated fairly highly with the interview response in the second half of his or her administration" (p. 55).

After coding the documents for each time period, Rasler and her colleagues rank ordered ten leaders for each of the six personality variables employed in Hermann's study. If temporal stability for each of the traits were found, then each leader would maintain their rank between time periods for each trait. Their analysis showed the highest correlation to be the nationalism trait (in-group bias), with a Spearman Rho correlation of .770 and the lowest to be the suspiciousness (distrust of others) with a Spearman Rho correlation of .297. Rasler and her

colleagues concluded that none of these were significantly correlated to warrant the view that these traits were in fact stable over time. In their conclusion, the authors suggested that the setting in which interviews took place (during conflict versus peacetime) and the source of the interview were potential explanations for the discrepancy in scores across periods.

A rarity in most cases with academic research, Hermann had the opportunity to respond to the claims made by Rasler, Thompson and Chester in the same issue of *International Studies Quarterly* (1980b). In her response, Hermann noted that Rasler, Thompson and Chester had neglected to take into account certain factors in their analysis. Most notable among the exceptions was Hermann's discussion of the debate between psychologists regarding temporal stability of personality traits. Hermann discussed previous research by Henry Alker who argued that situational variation fails to exclude personality as a useful explanatory tool (1972). This approach views individual personality as a critical component of the policy process, but at the same time understands personality to be at least somewhat situationally dependent. The most recent manual for Hermann's personality assessment takes this approach, stating:

"... the research on personality suggests that some people use contextual cues to determine what they do and, thus, may evidence changes in their trait scores depending on the nature of the situation. Other people's personalities are fairly stable across situations. By examining diverse material on a political leader, it is possible to determine how stable his or her leadership traits are." (Hermann 1999, p.36)

Rather than an either-or approach, the current view adopted by those who use the Leadership Trait Analysis technique is that certain settings provide opportunities for leaders to influence the policy process more than others. Differentiating between cases where the leader has more or less influence (whether it be in areas where the leader has prior policy experience or a personal interest in the policy's substance) is instrumental to understanding the role played by the

individual versus the role played by the context in which the individual is placed. Given this conception of personality, trait stability over time is not critical to the theory's validity. If stability is not found at all times, further research can be conducted to see whether leaders do show variation in cases where the context can be controlled.

In addition to the critiques made by Rasler and her colleagues, potentially contradictory findings exist within the research conducted by its proponents as well. The studies of the authorship effect conducted by Dille and Dille and Young discussed earlier point to a serious gap in our knowledge regarding the type of material that should be used in the creation of a trait analysis. While Dille and Young's research took a first cut at answering this question, they each used only two separate leaders in their analysis; used a limited sample of data to test the authorship hypothesis; and limited their analysis to only one of the seven traits studied by the Leadership Trait Analysis method.

A final critique is one that plagues any attempt at proving that individual decision-makers are important factors in the explanation and prediction of political behavior. It returns us to the very assumption of the neo-realist enterprise: that the distribution of power explains all that is important in international affairs, leaders may act differently and that may correspond with the behavior their states engage in - but this is simply a product of the environment in which they are placed. Switch George Bush for Saddam Hussein and each leader would act according to his surrounding environment and the power constraints of their new states. You may be able to measure personality scores, but those measurements are simply a product of the environment in which a leader is placed. If one were to examine the personality scores of a leader during a conflict and the period immediately following that conflict, you would find that the trait scores identified by Leadership Trait Analysis would shift accordingly.

Value of Alternative Approaches

While the above discussion points out the advantages of a decision making approach to foreign policy analysis and international relations, a brief cautionary note is necessary here. It is important to recognize that dogmatically adhering to any single approach to the understanding of human (or state) behavior is a sure path to flawed analysis. While our level of understanding is, regrettably, still weak enough that we must first identify how individual pieces of the puzzle help us explain what occurs in the world, it is rare that the behavior we are interested in better understanding is simply a function of any single cause. Rather than viewing alternative approaches as obstacles to our understanding of such behavior, in many cases competing approaches can actually be complementary and serve to better our understanding of events we are trying to explain. Looking at the Vietnam war we can see a whole host of both competing and complementary explanations for the United States' entanglement in that conflict.

By its nature as a military conflict, many explanations focus on the United States' military response to the North Vietnamese. The military tactics employed by the United States were directed at an enemy who was similarly equipped and failed to take into consideration the nature of guerilla warfare. While cutting supply lines to an armored division of tanks made sense when fighting Germany in World War II, supply lines in Vietnam consisted of bicycle paths and underground tunnels. While a conventional army relies on the support of a large military organization, guerrilla fighters can move from village to village when they have the support of the masses. In order to defeat the North Vietnamese, the United States would have had to completely decimate the population to destroy the support base for the guerrilla fighters.

While this perspective fits with the conventional wisdom thirty years after the war's end, during the war's prosecution many American lives were lost without a change in tactics. What can explain such seemingly irrational behavior? While the neo-realist approach focuses on the

broad palette of international politics and could be used to predict that proxy wars would occur between the Soviets and the Americans, we must look somewhere else to find explanations regarding the tactics employed by the United States' military. The bureaucratic politics literature in political science provides a clear reason for why the United States employed what now appears to be futile tactics when fighting the North Vietnamese. Institutions such as the United States military last much longer than the individuals who compose the organization. In order to preserve the knowledge that individuals possess and provide guidance for new members of the organization, standard operating procedures are used for both training and execution of the organization's goals. Prior to the Vietnam war, the armed forces of the United States trained for large scale wars to be fought against other great powers. The tactics necessary to defeat a small and highly motivated guerrilla army had not been developed at the time. In order for the organization to function effectively, the chain of command and standard operating procedures needed to be followed. Doing this, however, led to the use of strategies and tactics which virtually guaranteed defeat for the United States.

Another explanation for the United States' defeat in Vietnam moves closer to the systems level favored by neo-realists by focusing on the nature of the South Vietnamese government, rather than looking at specific leaders or entire organizations within the state. While the North was motivated under the leadership of Ho Chi Minh, many in the South viewed the regime of Ngo Dinh Diem as both brutal and corrupt (McMaster 1997, pp. 34-41). The lack of support for the South Vietnamese regime made defeating the highly motivated North a difficult task, one that could not be accomplished even with the help of the United States.

In addition to competing perspectives which focus on an alternative perspective, many complementary explanations for events in world politics exist which can be integrated to form a

more complete explanation of the phenomena political scientists attempt to understand. Continuing with the example of the United States war in Vietnam, we can see how individual level explanations can merge with bureaucratic explanations. H.R. McMaster's book, Dereliction of Duty: Lyndon Johnson, Robert McNamara, the Joint Chiefs of Staff, and the Lies that Led to Vietnam argues that the assignment of Robert McNamara as Secretary of Defense led to a deterioration of the relationship between the armed forces and the executive branch. McMaster argues that the coup which led to Diem's death was bitterly opposed by the Chiefs of Staff. While the coup was seen as a breakthrough by some in the executive branch, it was perceived quite differently elsewhere in the government.

“For his part McGeorge Bundy was relieved that the coup had finally been accomplished. He told the president and his closest advisers, “Now we have nothing to divide us.” Bundy's comment was unrealistically sanguine. Instead of fostering unity among agencies and individuals in the administration, the legacy of the Diem coup and assassination would inspire divisiveness and animosity. ... Responding effectively to a deteriorating situation in South Vietnam, furthermore, would require close coordination between civilian and military officials. Rather than relieving tensions between the [Joint Chiefs of Staff] and the Kennedy Administration, the “Asian Bay of Pigs” exacerbated them and made cooperation even more difficult” (McMaster 1997, p. 46)

From this quote we can see individual level factors (the perceptions held by McGeorge Bundy), institutional factors (the interplay between agencies within the United States government and military), and state level factors (the nature of the South Vietnamese regime leading to the violent coup).

While this study performs a review of the assumptions underlying a single technique within the individual level of analysis, it should not be taken as a rejection of alternative approaches. While narrow research questions may be better explained using one approach over another, in many cases those same questions will be best explained by integrating various

approaches to produce a more complete picture of the events in question. Moving to broader research questions (such as the reasons for United States involvement in the Vietnam war) demand the integration of approaches rather than blindly following a single factor's explanation of one component of the event.

Conclusion

While much has been done to develop personality assessment as an important tool in the study of foreign policy, we are still faced with significant questions regarding the legitimacy of this approach. The question asked in the title of this study, "State, Trait, or Design?" summarizes some of the critical issues the approach has yet to resolve. State versus trait assesses whether leaders possess stable and measurable personality traits or if they simply move from one characteristic to another depending on the political winds. The remaining component, design, asks whether the approach displays adequate sensitivity to distinguish between the instrumental speech so often attributed to our political leaders. If Leadership Trait Analysis is to be an effective tool for the analysis of foreign policy, more needs to be done to resolve these issues. This study attempts to contribute to the resolution of these issues through empirical tests of some of the key components of the technique: 1) the amount of data required to produce a reliable profile; 2) use of spontaneous versus prepared documents in profile creation; 3) trait stability over time, and 4) the degree to which the content of responses affects the trait scores produced by the system.

Hermann's manual for Leadership Trait Assessment recommends the use of at least 50 responses (of at least 100 words in length) in the production of a reliable personality profile. Yet, the physical limitations of hand coding documents have precluded studies to confirm that this

number is a satisfactory threshold. Using the public opinion strategy of sampling 1200 to 1500 respondents to arrive at +/- 3% of the actual public opinion, fifty responses seem inadequate. However, if we shift the unit of analysis to words then we would have at least 5000 items sampled in the fifty-response sample. Further analysis is necessary to test the validity of such samples.

Another issue which to date has been answered only in part is the use of prepared documents in profile creation. As noted above, Hermann excludes the use of non-spontaneous utterances in her work, arguing that such material is suspect and most likely does not reflect the leader who we are interested in studying. Rather, such statements are more likely reflective of those who have crafted the speech and suffer from a ghostwriting effect. Since we are interested in the leader's personality and not that of his or her speechwriter, Hermann limits her studies to spontaneous utterances. Unfortunately for scholars, many world leaders engage in interviews or press conferences on a limited basis. In cases of terrorist leaders such as Osama bin Laden, it is impossible to acquire spontaneous materials. With the exception of the Dille and Young (2000) study on conceptual complexity, we do not have rigorous studies that test the validity of the claim that only spontaneous utterances should be used.

The question raised by Rasler and her colleagues has also been left with an incomplete response. Because of the limited amount of data points used in previous studies, questions regarding the stability of the seven personality traits measured in Leadership Trait Analysis are still valid. The labor involved in the hand coding of documents has thus far prevented a thorough and rigorous analysis of trait score stability over time for all seven traits for a number of individuals. While we are still uncertain as to whether Rasler, Thompson and Chester's critique was accurate, the issue they raised is undoubtedly significant and is worthy of further attention.

A final issue that previous studies in Leadership Trait Assessment have not addressed is the situation in which the leader is placed. As Snyder, Bruck and Sapin pointed out, not only is the perceptions or characteristics of the leader important, but the leader's surroundings are also an important part of the decisions they make. As all leaders are not the same and thus make different decisions, each situation differs across leaders, whether it be within a country or between different countries. The leader of a war-torn Balkan nation will surely find themselves in a different setting than the leader of a country that has been at peace with its neighbors for the last hundred years. If we believe the leader's *past* experiences and environment influence their current decision-making, should we not also assess whether their *current* environment acts an influence on their behavior?

Each of these issues will be examined in the remaining chapters, specifically, the following questions will be addressed:

- Should non-spontaneous utterances be excluded from the material used to create profiles?
- Do the seven traits examined in Leadership Trait Analysis remain stable across time?
- Are the trait scores reflective of the leader's personality, or simply a function of their environment?
- What constitutes an appropriate sample size for the production of a reliable profile?

Chapter 2

Research Design

The inspiration behind the questions asked in this study are a product of countless hours of hand coding press conferences and interviews using the Leadership Trait Analysis technique. As for whether or not spontaneous material best reflects the leaders “true” personality, the logic behind the decision to use such material made sense but the lack of empirical tests using substantial amounts of data brought my skeptical nature to the forefront. While I am fairly confident that the spontaneous material best reflects the leader’s personality, I am much less certain regarding the answers to the other questions in this dissertation. The one thing which there is little doubt is that previous attempts at answering such questions were severely limited by the hand coding process. Whether it be a function of small sample sizes, human coding fatigue or subjectivity in the coding process, hand coding of material makes answering such questions with suitable confidence virtually impossible.

While the inspiration behind this study stems from hand coding, the ability to conduct the analysis found within these pages is a product of a rare opportunity to be directly involved with the people who are making automated coding a possibility. For the last two years I have worked as an analyst with Social Science Automation, Inc. (the developer of the Profiler Plus content analysis software used in this study). My position with Social Science Automation has allowed me an invaluable opportunity to test the questions that had remained in the back of my mind for the last few years. I was able to not only work on the continued development of the Leadership Trait Analysis coding scheme, but interact with Margaret Hermann (the developer of the technique) and was given access to data for a population of world leaders to be used as a

comparison base for the presidents in this analysis. Social Science Automation was generous in allowing me access to such data and the use of company resources to code the vast amounts of material used in this analysis. The remainder of this chapter outlines the specific research design employed in this dissertation.

After reviewing the literature regarding the use of Leadership Trait Analysis in the study of foreign policy decision-making, a number of issues can be identified which have either not been addressed or would benefit from additional research.

- Should non-spontaneous utterances be excluded from the material used to create profiles?
- What is an appropriate sample size for producing reliable profiles?
- Do the seven traits examined in Leadership Trait Analysis remain stable across time?
- Do trait scores reflect leader personality or are they simply a function of the topics the leader talks about in speeches and interviews?

Each of these questions will be addressed by performing the Leadership Trait Analysis method on all press conferences, interviews and spoken remarks given by four modern U.S. Presidents during their first four years in office. These documents were collected from the Public Papers of the Presidents series on CD-ROM, except for the most recent President in the study whose documents were gathered from a White House e-mail delivery system. Modern U.S. Presidents were chosen for testing the questions above for the following reasons. 1) Unlike many world leaders, U.S. Presidents provide scholars with a vast amount of publicly available English language texts from which we can construct personality profiles. 2) Selecting four modern U.S. Presidents also controls for a number of variables that might otherwise explain variation in our

dependent variables. Culture, language community, organizational role and a host of other factors all remain relatively consistent among U.S. Presidents⁶.

The specific Presidents chosen for this study are Presidents Truman, Johnson, Bush (senior) and Clinton. Previous research using Leadership Trait Analysis (Preston 1996, 2001) have found that both Truman and Johnson score low on the conceptual complexity trait when compared to other U.S. Presidents and high on the need for power trait. Presidents Bush and Clinton score in an opposite manner. On conceptual complexity both Presidents score higher than most other Presidents and on need for power they score lower. These pairings of high-low scores between each of the two sets of Presidents will provide interesting cases when testing the hypotheses found in Chapters 4 and 5 (described in more detail below) which assess trait stability across response topic and over for each of the leaders included in this analysis.

Chapter 3 will address what type of documents (spontaneous versus prepared) should be used in the construction of the profile. Both statistically and substantively significant differences for trait scores are expected based on arguments regarding speechwriting and attempts at sending a message to an audience (Herman, 1983, 1999). This will be tested in the first hypothesis:

Hypothesis 1: Each President's trait scores derived from prepared utterances will vary significantly from trait scores derived from spontaneous utterances.

This hypothesis will be tested by performing an ANOVA test of each personality trait score to determine whether those scores vary significantly by level of spontaneity. If the test fails to show significant variation by spontaneity type, Hypothesis 2 will be tested.

Hypothesis 2: Differences in trait scores derived from prepared versus spontaneous

⁶ While there are likely as many definitions of Political Culture as there are of personality, this study employs a conception of Culture as presented in Valerie Hudson's *Culture and Foreign Policy* (1997). Hudson identifies culture as shared meanings and value preferences. While these factors are by no means universal within the United States (or across time as is the case with the four Presidents included in this study), when compared to the population of world leaders as a whole, these factors remain relatively consistent for the leaders included in this analysis.

utterances are random.

Because speeches and remarks have been prepared in advance (and most likely written either completely or in part by others) they are not expected to accurately reflect the “real” personality of the leader in question. Whether it is the influence of other writers or an attempt at instrumental speech, the trait scores measured by the coding scheme should vary. These differences are expected to be random because of varying degrees of influence on the part of speechwriters as well as the specific words chosen to convey a desired message in prepared material. If the differences in scores are found to occur randomly or do not track in a predictable fashion, we will remove prepared utterances from our dataset for testing the remaining hypotheses. If there is a predictable shift in trait scores between spontaneity type, the prepared material may be included in the dataset with scores modified to adjust for this shift. The final hypothesis to be tested in this chapter will examine whether the scores used in the remaining analysis show significant variation by President.

Hypothesis 3: President is a significant source of variation in trait scores across all data coded in this study.

If the results fail to show statistically significant difference across each of the Presidents, the technique will not be a useful method for determining how the individual differences of leaders affect how they behave in the international arena.

Chapter 4 will examine forces external to the leader, assessing whether the trait scores are a) a result of the topical content of the data used to produce a profile; b) representative of the leader’s personality; c) or some combination of both. Examining the topical content of the leader’s speeches and interviews will assess whether scores reflect the topics they discuss. Each response to an interview or press conference question, as well as each topical section of a speech will be given a response code to place it into one of the following categories: conflict/security,

economic, political, or “other”. ANOVA analysis will then be performed to test whether response topic serves as a significant source of variation in trait scores.

Hypothesis 4: Mean trait scores for each President will exhibit statistically significant variation across response topic.

If statistically significant variation in mean trait scores is found across response topic, Hypothesis 5 will be tested to determine whether the variation is substantively important.

Hypothesis 5: Trait scores which exhibit statistically significant variation based on response topic will also exhibit substantively significant variation.

If both statistically significant and substantive variation is found based on response topic, then Hypotheses 6a-c, ANOVA tests for difference by leader, will be performed separately for each response topic as was done in Hypothesis 3 to ensure that differences can be identified across leaders. If we fail to identify differences between the four leaders for any trait-topic combination, the differences in scores by leader found in Chapter 3 may be a function of the frequency of different response topics in a leader’s collection of statements.

Hypothesis 6a: President is a significant source of variation in trait scores across all conflict/security responses coded in this study.

Hypothesis 6b: President is a significant source of variation in trait scores across all economic responses coded in this study.

Hypothesis 6c: President is a significant source of variation in trait scores across all political responses coded in this study.

If differences are found, trait stability across time will be tested separately for each topic in Chapter 5.

Chapter five revisits the debate between Hermann and Rasler et al (1980) regarding the temporal stability of the personality trait scores derived by Leadership Trait Assessment.

Drawing from the debate between Hermann and Rasler et. al., Chapter 5 tests for temporal stability of personality traits in two ways, the first is tested through Hypothesis 7.

Hypothesis 7: Mean trait scores for each time period will be classified in the same category ranking⁷.

The current method for determining whether a leader is high, low, or moderate on any given trait is to assess whether their mean trait score is within one standard deviation of the mean trait score for the entire population. If the leader's trait score falls within +/- one standard deviation of the population mean they are considered Moderate; if the score is one standard deviation above the population mean or more they are considered High; and if the score is one standard deviation below the population mean or more they are considered Low. Using this method for production of the range of accepted variation will ensure that a leader's trait score does not shift more than what is used for standard comparison purposes.

The second method by which we can assess temporal stability is tested in Hypothesis 8.

Hypothesis 8: A President's rank for any given trait (compared to the other three Presidents in the study) will remain constant across the entire four years.

Where Hypothesis 7 focuses on a discrete range where the leader's trait scores may fluctuate, Hypothesis 8 is concerned with relative positions. Previous research has used year as the unit of aggregation for testing temporal stability (Hermann 1980, Rasler et al 1980, Dille 2000, Dille and Young 2000). Using the year as a unit of aggregation, however, increases the likelihood that stability will be found by regressing trait scores in each period toward the overall mean. In order to assess what effect aggregation unit has on stability, Chapter 5 will test trait score stability at three different units of aggregation: the year, the quarter and the day. Using Preston (1996, 2001) as a guide, we would assume Presidents Bush and Clinton's conceptual complexity scores should remain higher than Truman and Johnson's scores for each of the sixteen time periods.

⁷ The dataset of world leader trait scores was provided by Social Science Automation for this study.

Chapter 6 will examine an issue which was of significant concern before the automation of the Leadership Trait Analysis coding technique, and in some cases is still an issue – sample size. As discussed in the previous chapter, Hermann (1999) argues that, at a minimum, 50 spontaneous utterances drawn from a variety of contexts and topics must be used to create a reliable profile (although Hermann recommends using more responses if the data is available). Automation allows us to create profile scores based on *all* available documents and provides an opportunity to assess the reliability of such samples. While automation of the content analysis process renders sample sizes a moot point for leaders with many available documents, gathering even 10 spontaneous documents for some world leaders is a difficult task. For such leaders, a determination of what constitutes a reliable sample size is necessary to adequately assess the reliability of profiles based on limited statements.

While public opinion polling and other surveys base sample size creation on a closed population (the population of the United States is a relatively constant figure), a sample of public statements made by leaders is based on an open population – one with no finite boundaries. This population has no definable size limit and its characteristics can vary from one leader to another. The thoughts, both expressed and unexpressed, in any individual's mind is infinite. In addition, what percentage of different statement topics, audience type, and environmental contexts (the combinations are endless) are appropriate to produce a representative sample?

Hypothesis 9: Profile scores based on 50 responses constitute a reliable sample of material to construct a personality profile.

Comparing the profile scores derived from all available documents to randomly selected samples of varying sizes will test Hypothesis 2a. If a threshold is found where a sample of 'x' responses

produces mean trait scores that fall within +/- 5% of the mean scores produced by the entire sample of data an additional test will be conducted. If, however, the tests in Chapter 5 fail to show trait stability over time, this test will become unnecessary. If trait scores fluctuate significantly from one time period to the next it leaves us without a single value by which to assess the accuracy of samples used for approximation. More importantly, if scores fluctuate across time or topic, limiting an analysis to a sample of all available material will overlook important shifts in trait scores. If this is found to be true, Chapter 6 will perform a brief test of 10 randomly generated samples of 50 responses to illustrate the consequences of attempting to approximate the value of a moving target.

Chapter 3

Design or Reality? Differences in Trait Scores by Spontaneity

Ask the average U.S. citizen how often politicians tell the truth, and you are very likely to receive an emphatic “never!” as a response. When dealing with foreign policy in general and matters of national security, we may even wish that our leaders refrain from speaking the truth to prevent an adversary from gathering critical knowledge of our strengths and weaknesses. Given this skepticism that our leaders are telling the truth, how can we be sure that what we measure when using techniques such as the Leadership Trait Analysis coding system is anything but a well crafted message the politician’s advisors have designed in their offices? This chapter assesses whether the Leadership Trait Analysis technique can differentiate between the personality the leader wants the public to see and their “true” personality.

In developing a system to measure the personality of such individuals, Hermann made a number of choices about how to best measure a leader’s personality. Hermann advocates the use of spontaneous utterances (non-prepared responses during interviews and press conferences) to prevent any ghost writing effect from biasing the results of the assessment. While some leaders spend more time than others creating the speeches they deliver at various events, most world leaders have a team of speechwriters producing the bulk of the words they speak at public events. By using only spontaneous utterances found in press conferences and interviews, the words spoken by the leader are presumed to be more representative of the leader’s personality than those written by their advisers.

While Hermann’s decision to exclude prepared material intuitively makes sense, to date no empirical tests have been conducted on all seven traits to assess whether a difference exists between spontaneous and prepared material for world leaders. Two studies have examined

whether scores for particular personality traits produced from spontaneous material differ from those scores produced strictly from prepared material. Brian Dille (2000) measured the effect of spontaneity for Operational Code indices and conceptual complexity by examining speeches and press conference materials for two U.S. Presidents: Ronald Reagan and George Bush, Sr. Using a sample size of thirty documents, he found significant variation between source type (spontaneous versus prepared utterances) for the Operational Code indices, but found no significant variation for conceptual complexity.

In contrast, Dille and Michael Young (2000) examined the authorship question for two other U.S. Presidents, Jimmy Carter and Bill Clinton. In this study Dille and Young did find significant variation in conceptual complexity scores based on source type. In fact the differences between the scores in each case were of almost identical magnitude, leading the authors to conclude that the difference in scores may be a function of the style and setting of press conferences versus the delivery of prepared remarks. “It may be that the higher variety of issues raised in press conferences or other question-and-answer sessions naturally raises the conceptual complexity of the individual answering the multiple queries” (p. 593). The other possibility presented by Dille and Young is that the speechwriters craft the nature of the speech to reduce the level of complexity of the speaker in order to conform to some designated acceptable level for the public.

A number of factors serve as plausible explanations for the conflicting results found in these studies. In addition to the suggestions made by Dille and Young, at least three other explanations can be identified to explain these conflicting results. First, conceptual complexity scores as measured in these studies may be a function of the sample size used in each analysis. In the Dille study, the total number of documents used to produce trait scores was thirty. While this satisfies

the minimum requirements suggested by Hermann in the coding manual for Leadership Trait Analysis, to date no empirical studies have been conducted to assess what number of documents is sufficient to produce a reliable trait analysis. In their study of Presidents Carter and Clinton, Dille and Young's results were based on more documents (approximately 80 prepared and 80 spontaneous for each President), yet still consisted of a sample of the total population of statements made by the two leaders.

A second explanation may simply be that spontaneity of statements is not a relevant factor for some leaders. If the differences in complexity scores found by Dille and Young were a product of instrumental speech, it may be that some Presidents find little use for instrumentality and their prepared remarks reflect what they would say in spontaneous settings. On the other hand, some Presidents may be skillful enough in their use of instrumental speech to mask any difference in their statements (at least enough to prevent the Leadership Trait Analysis technique from identifying the difference). A final explanation for the conflicting results may lie with the scope of the studies in question. Because these studies have only examined differences by authorship for conceptual complexity, we are missing information on the six remaining personality traits. There may be a broader pattern either across traits or across leaders which can be identified by testing for differences between spontaneous and prepared material for the remaining six traits.

This chapter addresses these issues by testing the following hypotheses:

Hypothesis 1: Each President's trait scores derived from prepared utterances will vary significantly from trait scores derived from spontaneous utterances.

This hypothesis will be tested by performing an ANOVA test of each personality trait score to determine whether those scores vary significantly by level of spontaneity. If the test fails to show

significant variation by spontaneity type, Hypothesis 2 will be tested.

Hypothesis 2: Differences in trait scores derived from prepared versus spontaneous utterances are random.

Because speeches and remarks have been prepared in advance (and most likely written either completely or in part by others) they are not expected to accurately reflect the “real” personality of the leader in question. Whether it is the influence of other writers or an attempt at instrumental speech, the trait scores measured by the coding scheme should vary. These differences are expected to be random because of varying degrees of influence on the part of speechwriters as well as the specific words chosen to convey a desired message in prepared material. If the differences in scores are found to occur randomly or do not track in a predictable fashion, we will remove prepared utterances from our dataset for testing the remaining hypotheses. If there is a predictable shift in trait scores between spontaneity and prepared material, the prepared material may be included in the dataset with scores modified to adjust for this shift.

Data used in this study

Studies using the Leadership Trait Analysis technique previously faced enormous labor costs associated with hand coding documents for each of the seven personality traits measured by the technique. Hermann identified a minimum threshold of fifty responses of one hundred words in length spread across a variety of topics and drawn from as many separate events as possible. Taking these criteria as a guide, we can identify the following requirements as the minimum acceptable for a reliable profile:

Summary 3.1 Minimum Sample Requirements (taken from Hermann 1983,1999)

Number of Distinct Documents: no minimum given

Number of Responses: 50

Number of Words: 5,000

Because of the immense labor costs and reliability issues associated with hand coding, most studies conducted using the Leadership Trait Analysis technique were performed by a single researcher and used little more than the minimum requirements listed above. With the advent of automated content analysis software, coding much more than the minimum suggested amount of data poses little additional cost when compared to hand coding the same material. With automated coding schemes the only costs associated with coding additional documents are related to collection and preparation of the material⁸. In order to be coded by the Profiler Plus content analysis software, documents must be in electronic format (in the standard Windows .txt format). The only required preparation of the documents after they are in electronic format consists of marking text that was not spoken by the leader in question to be “ignored” (so the coding software will skip that material when coding). Depending on a researcher’s interests, additional document mark-up may be performed which will add additional preparation time to the process. For the present study, document codes were inserted into each document that identifies the date in which the speech or press conference was given and whether the document was spontaneous or prepared. In cases of press conferences where a prepared statement usually precedes the question and answer session, each section was given its own code (because of this the total number of actual events will be slightly less than the number of prepared and spontaneous documents combined).

Because there have been no tests to verify that the minimum response and word requirements suggested by Hermann produce reliable trait scores, this study uses all available documents to reduce the likelihood that trait scores are an artifact of the specific responses used

⁸ Documents for this study were collected from the Public Papers of the Presidents series on CD-ROM, with the exception of President Clinton’s documents. Clinton’s speeches, press conferences, and interviews were collected by Social Science Automation, Inc. from the Clinton Administration’s official White House Press Office list-serve on the Internet and were generously made available for use in this study.

in our analysis (see Chapter 7 for a test of acceptable sample sizes). The information in Summaries 3.2 through 3.5 illustrates the amount of data used in this study and how it compares to the minimum data requirements as suggested by Hermann.

Summary 3.2 Document Information for George Bush, Sr.

Total Documents: 2,184
 Total Responses: 16,889
 Total Responses over 100 words: 10,261
 Total Words: 3,286,643

Prepared Material

Documents: 1,673
 Responses: 9,178
 Responses over 100 words: 7,000
 Words: 2,413,265

Spontaneous Material

Documents: 511
 Responses: 7,711
 Responses over 100 words: 3,261
 Words: 873,378

Prepared Documents compared to Minimum Suggested Profile

180 x the Minimum Suggested Responses (not using 100 word minimum criteria)
 140 x the Minimum Suggested Responses (using 100 word minimum criteria)
 480 x the Minimum Suggested Words

Spontaneous Documents compared to Minimum Suggested Profile

150 x the Minimum Suggested Responses (not using 100 word minimum criteria)
 65 x the Minimum Suggested Responses (using 100 word minimum criteria)
 175 x the Minimum Suggested Words

Summary 3.3 Document Information for Bill Clinton

Total Documents: 1,380
 Total Responses: 13,014
 Total Responses over 100 words: 7,724
 Total Words: 2,409,498

Prepared Material

Documents: 1,002
 Responses: 8,518
 Responses over 100 words: 5,041
 Prepared Words: 1,732,573

Spontaneous Material

of Documents: 378
 # of Responses: 4,496
 # of Responses over 100 words: 2,683
 # of Words: 676,925

Prepared Documents compared to Minimum Suggested Profile

170 x the Minimum Suggested Responses (not using 100 word minimum criteria)
 100 x the Minimum Suggested Responses (using 100 word minimum criteria)
 345 x the Minimum Suggested Words

Spontaneous Documents compared to Minimum Suggested Profile

90 x the Minimum Suggested Responses (not using 100 word minimum criteria)

55 x the Minimum Suggested Responses (using 100 word minimum criteria)

135 x the Minimum Suggested Words

Summary 3.4 Document Information for Lyndon Johnson

Total Documents:	1,265
Total Responses:	11,923
Total Responses over 100 words:	7,375
Total Words:	1,783,675

Prepared Material

Documents:	1,143
Responses:	10,047
Responses over 100 words:	6,581
Prepared Words:	1,523,581

Spontaneous Material

Documents:	122
Responses:	1,876
Responses over 100 words:	794
Words:	260,094

Prepared Documents compared to Minimum Suggested Profile

200 x the Minimum Suggested Responses (not using 100 word minimum criteria)

130 x the Minimum Suggested Responses (using 100 word minimum criteria)

305 x the Minimum Suggested Words

Spontaneous Documents compared to Minimum Suggested Profile

40 x the Minimum Suggested Responses (not using 100 word minimum criteria)

15 x the Minimum Suggested Responses (using 100 word minimum criteria)

50 x the Minimum Suggested Words

Summary 3.5 Document Information for Harry Truman

Total Documents:	521
Total Responses:	7,162
Total Responses over 100 words:	929
Total Words:	772,229

Prepared Material

Documents:	374
Responses:	1,463
Responses over 100 words:	817
Words:	633,478

Spontaneous Material

Documents:	147
Responses:	5,699
Responses over 100 words:	112
Words:	138,751

Prepared Documents compared to Minimum Suggested Profile

30 x the Minimum Suggested Responses (not using 100 word minimum criteria)

15 x the Minimum Suggested Responses (using 100 word minimum criteria)

125 x the Minimum Suggested Words

Spontaneous Documents compared to Minimum Suggested Profile

115 x the Minimum Suggested Responses (not using 100 word minimum criteria)

2 x the Minimum Suggested Responses (using 100 word minimum criteria)
25 x the Minimum Suggested Words

Regardless of whether the data is limited to only spontaneous material or whether both spontaneous and prepared material is included, the documents coded in this study surpass by far the amount suggested by Hermann to produce a reliable profile. While the minimum suggested number of words is approximately 5,000 words, automated content analysis allows us to code over three million words for Bush (at the high end of our leaders) and over 750,000 words for Truman (at the low end). The suggested number of responses was also surpassed by over 7,000 for Truman and by nearly 17,000 for Bush. After restricting the data to only spontaneous material we are left with over 7,000 total responses for Bush (at the high end) and over 1,800 responses for Johnson (at the low end). These numbers include responses that have less than 100 words (the minimum suggested by Hermann). If we exclude the responses that have less than 100 words we are left with slightly more than 3,261 responses for Bush, 2,683 for Clinton, 794 for Johnson and only 112 for Truman.

While Hermann's original work was based only on responses of at least 100 words in length, this study will include all available responses regardless of word count. While this decision has the advantage of increasing the total number of responses in each leader's sample, there is a theoretical reason for choosing to include these responses. Take the following responses as an example: "Absolutely." "Without a doubt." "It depends." Each of these responses is useful in determining conceptual complexity scores. Brief statements from a leader who frequently uses such responses to reporter's questions are just as reflective of their level of conceptual complexity as responses that total 100 words or more. It can even be argued that such direct responses to a question by a reporter may be *more* reflective of a leader's level of conceptual complexity because they choose to speak as definitively as possible. This is

applicable for other traits as well. Consider the distrust and in-group bias traits. The following statement: “The United States is the most powerful country in the world and we will defeat the evil designs of the Soviet Union” while only twenty two words in length, would score as positive for both the distrust (evil designs of the Soviet Union) and the in-group bias (United States is the most powerful) traits. Limiting ourselves to responses of only 100 words or more in length will unnecessarily limit the material from which a profile can be created and, potentially, result in less reliable assessments.

Truman serves as an excellent example of why we should not restrict our data based on the number of words per response. Looking at the data collected and coded for Truman (see Summary 3.5) we can see that Truman’s responses to questions raised in interviews and press conferences were much shorter than any of the other Presidents included in this analysis. With only 112 of nearly 5,700 responses composed of at least 100 words, the vast majority of Truman’s responses would need to be excluded from this analysis. Given that the overwhelming majority of his responses are less than 100 words we can assume that brief (one might say “to the point”) responses are his natural way of responding to questions. Excluding these responses would thus lead to a biased assessment of his personality.

In the past, hand coding reduced the amount of data we were able to code for each trait. When limited to only 50 responses it was important to ensure that there was a significant amount of data to establish a profile. If 100 responses were used and many were less than 100 words the sample of data used to produce the profile would be severely limited. Now that the traditional costs associated with coding every publicly available document has been overwhelmingly reduced, we would undoubtedly be excluding valuable information when producing personality trait scores for leaders.

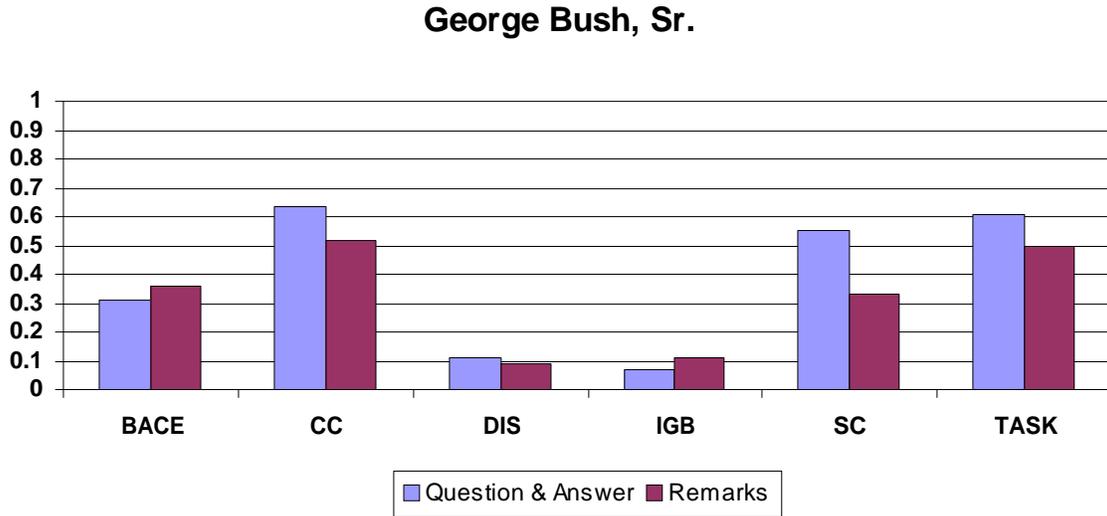
Findings

Table 3.1 and Figure 3.1 show the results of the ANOVA on spontaneity test for George Bush. As the data in Table 3.1 show, statistically significant variation was found for six of the seven personality traits for Bush. With the exception of Need for Power, each of Bush's trait scores varied significantly ($p < .0001$) based on type of utterance. An initial glance at Figure 3.1 shows not only statistical significance, but what appears to be substantively significant differences between scores based on spontaneity type for at least three of Bush's personality traits (conceptual complexity, self-confidence, and task orientation).

Table 3.1 Spontaneity ANOVA for George Bush, Sr.

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	.0001	Question & Answer	6,102	842,681	.31
		Remarks	8,051	2,302,759	.36
CC	.0001	Question & Answer	6,510	859,088	.63
		Remarks	8,965	2,401,272	.52
DIS	.0001	Question & Answer	5,212	796,556	.11
		Remarks	8,454	2,347,212	.09
IGB	.0001	Question & Answer	4,169	706,584	.07
		Remarks	7,392	2,217,437	.11
PWR	.10	Question & Answer	6,069	841,453	.25
		Remarks	7,977	2,293,680	.24
SC	.0001	Question & Answer	5,962	829,555	.55
		Remarks	6,963	2,141,665	.33
TASK	.0001	Question & Answer	5,790	828,188	.61
		Remarks	8,694	2,377,163	.49

Figure 3.1 Statistically Significant (.05 level) Trait Score Differences by Spontaneity for Bush



While Figure 3.1 illustrates what appears to be sizeable shifts in at least three of the mean trait scores, determining whether these shifts are substantively significant is a different issue altogether. The most meaningful way of assessing whether the observed differences in trait scores were substantively significant would be to match the differences to behavioral predictions derived from the Leadership Trait Analysis literature. Unfortunately, sophisticated correlations between trait scores and behavior have yet to be developed, preventing an assessment of the significance of the differences found between trait score means based on spontaneous and prepared material. Until such connections can be developed, we must refer to the literature on Leadership Trait Analysis to ascertain the substantive significance of the variation.

Hermann’s method of interpreting the scores produced by the technique involved comparisons with a broader population of leaders. If the mean trait score for a leader is one standard deviation above the population’s mean score or more for that trait, the leader is characterized as “High” on that trait. If the leader’s mean trait score is one standard deviation below the population’s mean score or less for that trait, the leader is characterized as “Low” on

that trait. Using the standard deviation of the population’s mean scores for each trait, a leader who is characterized as “High” on any given trait means they score higher on that trait than 84% of the leaders included in the population. Those who are characterized as “low” score lower than 84% of the leaders included in the population⁹.

Using this technique, we can compare the mean trait scores for both spontaneous and prepared material for each leader to a broader population of leaders to identify category ranks for each trait. Mean trait scores and standard deviations for a population of 51 world leaders was generously made available by Social Science Automation for comparison purposes. This information is presented in Table 3.5 below.

Table 3.2 Population Mean Trait Scores and Comparison Ranges

Trait	Population Mean	Standard Deviation	Low Range	Moderate Range	High Range
BACE	.35	.04	<= .31	>.31 and <.39	>= .39
CC	.57	.04	<= .53	>.53 and <.61	>= .61
DIS	.12	.04	<= .08	>.08 and <.16	>= .16
IGB	.09	.02	<= .07	>.07 and <.11	>= .11
PWR	.24	.03	<= .21	>.21 and <.27	>= .27
SC	.41	.09	<= .32	>.32 and <.50	>= .50
TASK	.63	.06	<= .57	>.57 and <.69	>= .69

Using the data in Table 3.2, we can classify Bush on each trait if only spontaneous or only prepared material was used in the calculation of mean trait scores. The results of this assessment are reported in Table 3.3. We find that not only is there statistically significant differences for six of the personality traits but substantively significant variation occurs on five of those six traits. With the exception of distrust of others, Bush would fall into different categories depending on

⁹ The calculation of population mean scores and standard deviations for each of the seven traits is calculated using each leader’s mean trait scores as opposed to a pooled set of all observations for all leaders. This use of “person” means and variances rather than all observations is taken from a recommendation included in a study conducted by Dr. Herbert Weisberg commissioned by Social Science Automation, Inc.

whether spontaneous or prepared material was used to calculate the trait scores. In three of the five traits where substantive change is observed, the change moves Bush across the entire comparison spectrum (either from high to low or low to high).

Table 3.3 Substantive Difference in Trait Scores by Spontaneity Type for George Bush, Sr.

Trait	Spontaneous Category	Prepared Category	Change?
BACE	Low	Moderate	Yes
CC	High	Low	Yes
<i>DIS</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
IGB	Low	High	Yes
<i>PWR *</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
SC	High	Low	Yes
TASK	Moderate	Low	Yes

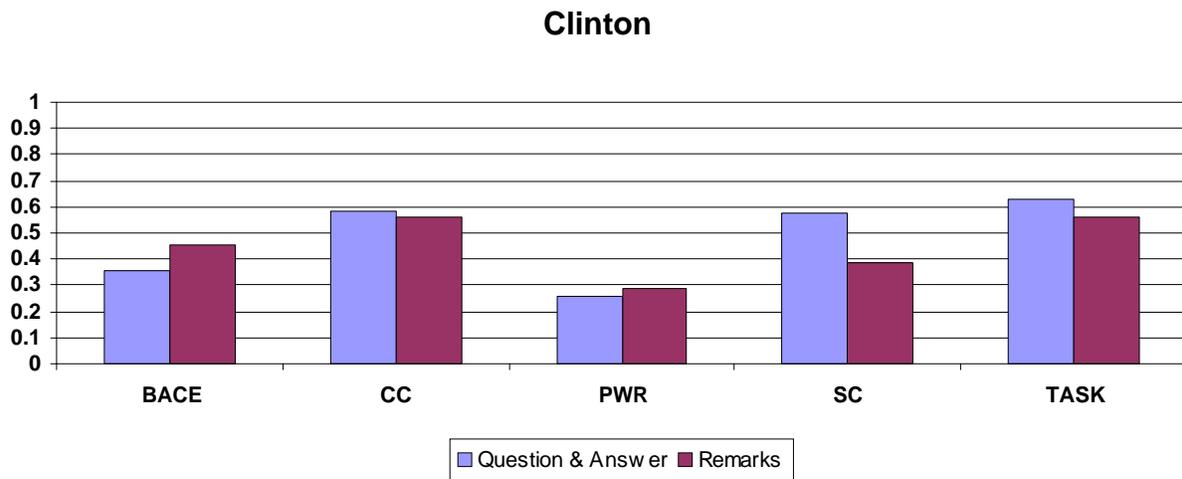
* No statistically significant difference found between spontaneous and prepared scores for this trait.

Examining the ANOVA test results for Bill Clinton we find similar results (see Table 3.4). Clinton also shows statistically significant variation on five of the seven traits (with in-group bias just exceeding the .05 level of confidence). While Bush's need for power failed to show statistically significant variation based on the spontaneity of the material, Clinton's distrust of others and in-group bias traits fail to show significant variation. In addition, an initial glance at Figure 3.2 shows less sizable differences in the trait scores based on spontaneous versus prepared material than was seen in Bush's material.

Table 3.4 Spontaneity ANOVA for Bill Clinton.

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	.0001	Question & Answer	3,963	666,089	.36
		Remarks	7,351	1,642,222	.45
CC	.0001	Question & Answer	4,008	669,191	.59
		Remarks	8,166	1,715,503	.56
DIS	.20	Question & Answer	3,385	626,886	.09
		Remarks	7,261	1,633,161	.10
IGB	.0515	Question & Answer	3,206	606,041	.06
		Remarks	7,116	1,607,774	.07
PWR	.0001	Question & Answer	3,930	664,950	.26
		Remarks	7,227	1,631,488	.29
SC	.0001	Question & Answer	3,742	642,116	.58
		Remarks	5,630	1,449,649	.38
TASK	.0001	Question & Answer	3,767	660,447	.63
		Remarks	7,928	1,696,161	.56

Figure 3.2 Statistically Significant (.05 level) Trait Score Differences by Spontaneity for Clinton



While the initial glance at Figure 3.2 leads us to believe that there was little substantive variation between Clinton’s trait scores based on spontaneity, using the population of world leaders as a comparison base we find that Clinton would be placed in different categories for four

of the five significantly different traits depending on which material was used to calculate trait scores (see Table 3.5 for details). Belief in one’s ability to control event, need for power, self-confidence and task orientation all shift from one category to another based on the spontaneity of the material used to establish a profile score. Of the traits that showed statistically significant variation based on spontaneity, conceptual complexity shows no substantive change between document type (as opposed to the shifts identified by Dille and Young in their sample of 160 documents). While in-group bias narrowly missed the .05 level of confidence, it did show substantive variation when compared to the population data used for categorization. While some of Bush’s trait score shifts moved across the entire continuum of high to low, each of Clinton’s trait score shifts were limited to moving towards or away from a moderate classification.

Table 3.5 Substantive Difference in Trait Scores by Spontaneity Type for Bill Clinton

Trait	Spontaneous Category	Prepared Category	Change?
BACE	Moderate	High	Yes
CC	Moderate	Moderate	No
<i>DIS *</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
<i>IGB *</i>	<i>Low</i>	<i>Moderate</i>	<i>Yes</i>
PWR	Moderate	High	Yes
SC	High	Moderate	Yes
TASK	Moderate	Low	Yes

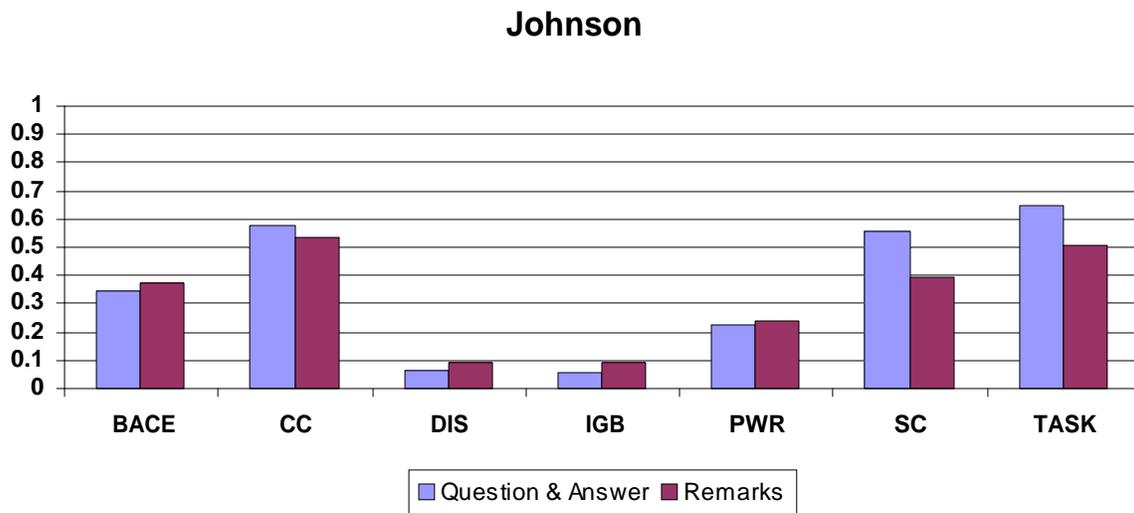
* No statistically significant difference found between spontaneous and prepared scores for these traits.

Performing these same tests for Johnson’s personality traits, we find that all of his traits show highly significant variation ($p < .0005$ for one trait and $p < .0001$ for the others) based on the spontaneity of the documents used to produce trait scores (see Table 3.6 for details). An initial glance at Johnson’s trait score differences in Figure 3.3 shows what appear to be sizable substantive differences for a number of traits.

Table 3.6 Spontaneity ANOVA for Lyndon Johnson.

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	.0005	Question & Answer	1,661	257,199	.34
		Remarks	8,980	1,433,322	.37
CC	.0001	Question & Answer	1,570	256,095	.58
		Remarks	9,569	1,497,563	.54
DIS	.0001	Question & Answer	1,139	229,340	.06
		Remarks	8,130	1,348,731	.09
IGB	.0001	Question & Answer	1,134	229,226	.06
		Remarks	7,991	1,320,787	.09
PWR	.0001	Question & Answer	1,650	256,952	.22
		Remarks	8,901	1,425,258	.24
SC	.0001	Question & Answer	1,524	244,868	.56
		Remarks	6,864	1,182,530	.39
TASK	.0001	Question & Answer	1,425	249,885	.65
		Remarks	9,253	1,465,514	.51

Figure 3.3 Statistically Significant (.05 level) Trait score Differences by Spontaneity for Johnson



Taking a closer look at the data in Table 3.7 we find the graphical display to be quite telling. Of the seven statistically significant trait score differences between spontaneous and prepared material, four are also substantively significant. Distrust of others, in-group bias, self-

confidence and task orientation all show substantive variation depending on the spontaneity of the material, while Johnson's scores for belief in one's ability to control events, conceptual complexity and need for power all remain in the same category regardless of spontaneity. As was seen with Clinton, Johnson's substantive trait score shifts either moved towards or away from the moderate classification.

Table 3.7 Substantive Difference in Trait Scores based on Spontaneity Type for Lyndon Johnson

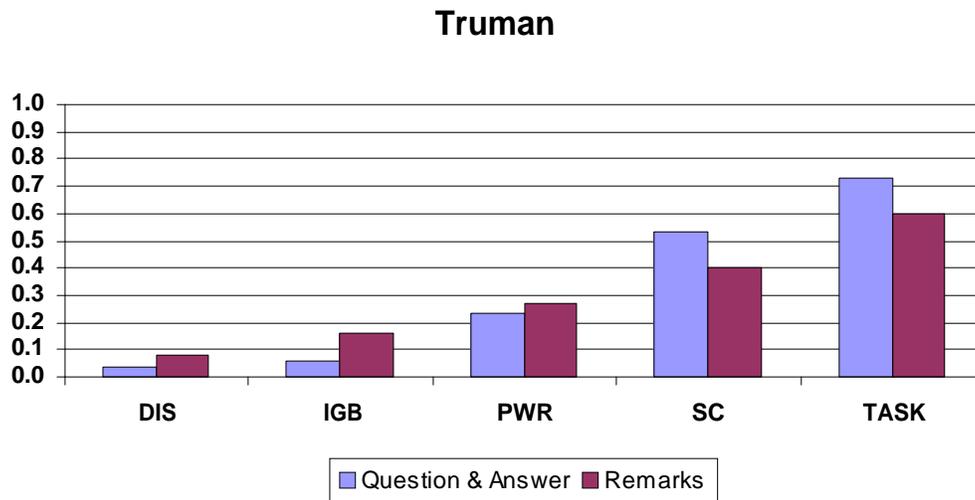
Trait	Spontaneous Category	Prepared Category	Change?
BACE	Moderate	Moderate	No
CC	Moderate	Moderate	No
DIS	Low	Moderate	Yes
IGB	Low	Moderate	Yes
PWR	Moderate	Moderate	No
SC	High	Moderate	Yes
TASK	Moderate	Low	Yes

Moving to Truman, we find similar results with variations in the significant traits. The data in Table 3.8 shows that five of the seven personality traits show statistically significant variation based on spontaneity. Distrust of others, in-group bias, need for power, self-confidence, and task orientation are all significant at the .001 level. Belief in one's ability to control events and conceptual complexity both fail to show statistically significant variation by spontaneity type. The initial glance at Figure 3.4 is similar to Johnson and Clinton in that some traits appear to show substantively significant differences.

Table 3.8 Spontaneity ANOVA for Harry Truman.

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	.2877	Question & Answer	3,404	110,047	.35
		Remarks	1,107	616,498	.37
CC	.1921	Question & Answer	3,236	110,224	.48
		Remarks	1,196	628,948	.50
DIS	.0001	Question & Answer	2,176	84,364	.04
		Remarks	1,095	624,365	.08
IGB	.0001	Question & Answer	733	40,005	.06
		Remarks	767	587,852	.16
PWR	.0010	Question & Answer	3,210	106,450	.23
		Remarks	1,085	615,101	.27
SC	.0001	Question & Answer	3,443	106,495	.53
		Remarks	957	589,241	.40
TASK	.0001	Question & Answer	2,219	84,648	.73
		Remarks	1,104	626,450	.60

Figure 3.4 Statistically Significant (.05 level) Trait score Differences by Spontaneity for Truman



Examination of the data in Table 3.9 shows that each of the traits that showed statistically significant differences in mean scores based on spontaneity type also show substantive differences. Distrust of others, in-group bias, need for power, self-confidence and task orientation all shift from one category to another depending upon the spontaneity type of

material used to produce mean trait scores. The range of Truman’s trait score shifts is identical to Clinton and Johnson’s in that they either move toward or away from the moderate category and never move across the entire continuum of high to low when compared to the population of leaders provided by Social Science Automation, Inc.

Table 3.9 Substantive Difference in Trait Scores by Spontaneity Type for Harry Truman

Trait	Spontaneous Category	Prepared Category	Change?
<i>BACE *</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
<i>CC *</i>	<i>Low</i>	<i>Low</i>	<i>No</i>
DIS	Low	Moderate	Yes
IGB	Low	High	Yes
PWR	Moderate	High	Yes
SC	High	Moderate	Yes
TASK	High	Moderate	Yes

* No statistically significant difference found between spontaneous and prepared scores for these traits.

Analysis

Hypothesis 1: Trait scores derived from prepared utterances will vary significantly from trait scores derived from spontaneous utterances.

The results from the ANOVA tests on spontaneity and trait score means partially confirm Hypothesis 1. Each President in this study had statistically significant variation based on the type of material used to produce trait scores on at least five of the seven personality traits used in Leadership Trait Analysis. Both the self-confidence and task orientations show statistically significant differences for all four Presidents while the remaining five traits showed statistically significant differences for three of the four leaders. Truman and Clinton had the least statistically significant trait totals at five traits, Bush showed statistically significant on six traits while Johnson showed statistically significant differences on all seven traits. In no cases did two leaders have the same trait appear statistically insignificant (see Table 3.10 for the full comparison of Presidents and statistically significant variation).

Table 3.10 Summary of Statistically Significant Variation on Spontaneity

Trait	Bush	Clinton	Johnson	Truman
BACE	Yes	Yes	Yes	No
CC	Yes	Yes	Yes	No
DIS	Yes	No	Yes	Yes
IGB	Yes	No	Yes	Yes
PWR	No	Yes	Yes	Yes
SC	Yes	Yes	Yes	Yes
TASK	Yes	Yes	Yes	Yes

Hypothesis 2: Differences in trait scores derived from prepared versus spontaneous utterances are random.

Examining the substantive implications of the variation in trait scores for the statistically significant variation we fail to confirm Hypothesis 2, that differences between prepared and spontaneous trait scores are random. Examining each trait individually by comparing the shifts for each President we can see that there is a definite pattern to the direction of the difference in scores based on spontaneous versus prepared material (See Table 3.11 for full information).

Table 3.11 Direction and Magnitude of Change between Trait Scores for Prepared and Spontaneous Material.

Trait	President	Direction of Change (Prepared to Spontaneous)	Magnitude
BACE	Bush	Negative	.05
	Clinton	Negative	.09
	Johnson	Negative	.03
CC	Bush	Positive	.11
	Clinton	Positive	.03
	Johnson	Positive	.04
DIS	Bush	Positive	.02
	Johnson	Negative	.03
	Truman	Negative	.04
IGB	Bush	Negative	.04
	Johnson	Negative	.03
	Truman	Negative	.10

Trait	President	Direction of Change (Prepared to Spontaneous)	Magnitude
PWR	Clinton	Negative	.03
	Johnson	Negative	.02
	Truman	Negative	.04
SC	Bush	Positive	.22
	Clinton	Positive	.20
	Johnson	Positive	.17
	Truman	Positive	.13
TASK	Bush	Positive	.12
	Clinton	Positive	.07
	Johnson	Positive	.14
	Truman	Positive	.13

For the statistically significant trait score differences, we do see a consistent pattern in terms of the direction of change. With the exception of distrust of others, the direction of change for each trait is uniform across leaders. Belief in one's ability to control events, in-group bias, and need for power all are lower in spontaneous utterances than they are in prepared remarks. Conceptual complexity, self confidence and task orientation all appear higher in spontaneous utterances than scores produced from prepared remarks. Distrust of others is the only trait that exhibits mixed results with Bush appearing more distrustful in spontaneous utterances while Johnson and Truman appear less distrustful in spontaneous utterances. This is especially interesting given the frequent criticism levied against psychological approaches to foreign policy analysis regarding instrumental speech. While it would be expected that spontaneous and prepared materials would differ based on instrumental speech it is by no means logical to assume that these differences would be in the same direction for all traits. In some cases a leader may wish to present a more conciliatory approach to an impending conflict and in others they may exaggerate their perception of hostility. Each of these approaches is likely to influence the scores produced for

distrust of others and in-group bias. Because different situations require different tactics we have little reason to believe that these differences based on spontaneity would be in the same direction for all leaders.

For other traits, such as conceptual complexity, we might expect to see the difference in a specified direction. In an analysis of Eisenhower's leadership style by Fred Greenstein (1995), Eisenhower is described as "[giving] vague, syntactically tangled answers to reporters' questions in press conferences, and his assertions that he was not familiar with matters that had been discussed in the New York Times" (p. 55). Greenstein argues that this was not a product of Eisenhower's incompetence as a leader, but reflective of his "hidden-hand leadership" style. Such answers would likely lead to a decrease in the conceptual complexity score of Eisenhower based on his spontaneous material. Interestingly enough, each of the Presidents who showed statistically significant variation on conceptual complexity in this study show the reverse to be the case. Presidents Bush, Clinton and Johnson all show significantly higher complexity scores in spontaneous settings than they do in prepared remarks.

While the direction may remain consistent for the leaders in all but the distrust trait, the magnitude of that change does not remain consistent. Not only does the magnitude change differ from leader to leader within traits, the magnitude change differs across traits for the same leader. Referring again to Table 3.11, we can see that Clinton's score difference for belief in one's ability to control events is greater than Bush and Johnson's scores. However, moving to conceptual complexity, we see that Clinton has the lowest magnitude shift of the three leaders with significant variation on that trait. This random variation in magnitude requires the exclusion of prepared material in the remaining questions asked by this study. While the data shows a distinct pattern (at least for the four Presidents included in this analysis) in the direction of

change between prepared and spontaneous materials, we are left without a clear pattern in the magnitude of those differences. Without a pattern it is impossible to develop a formula for adjusting prepared material to be used in further tests of the Leadership Trait Analysis technique.

As was discussed above, statistical significance is only meaningful if accompanied by substantively significant differences in trait scores between the two settings (see Tables 3.12a-g for a complete list of substantive shifts for statistically significant trait score differences). In the case of belief in one’s ability to control events, Presidents Bush and Clinton showed substantive variation between trait scores derived from prepared versus spontaneous material. In both cases the President switched from a higher comparison group to a lower comparison group when moving from prepared to spontaneous material. President Johnson also showed statistically significant variation between the two types of material, however he showed no substantive variation between the two types of material.

Table 3.12a Substantive Shifts for Belief in One’s Own Ability to Control Events

Trait	Pres	Spontaneous Category	Prepared Category	Change?
BACE	Bush	Low	Moderate	Yes
	Clinton	Moderate	High	Yes
	Johnson	Moderate	Moderate	No

For conceptual complexity we see a slightly different pattern with only Bush showing substantive change between categories, moving from the High to the Low category. Presidents Clinton and Johnson, who both showed statistically significant variation on this trait did not move from the Moderate category when compared to the population of world leaders used in this analysis.

Table 3.12b Substantive Shifts for Conceptual Complexity

Trait	Pres	Spontaneous Category	Prepared Category	Change?
CC	Bush	High	Low	Yes
	Clinton	Moderate	Moderate	No
	Johnson	Moderate	Moderate	No

In the case of distrust of others, the only trait which showed contradictory changes in direction for the four Presidents in this analysis, we only see one President moving from one category to another. Bush in fact moves from the lowest to the highest category for this trait, showing a much higher distrust of others in spontaneous settings than in his prepared remarks. Presidents Clinton and Johnson both possess moderate rankings for each category.

Table 3.12c Substantive Shifts for Distrust of Others

Trait	Pres	Spontaneous Category	Prepared Category	Change?
DIS	Bush	Moderate	Moderate	No
	Johnson	Low	Moderate	Yes
	Truman	Low	Moderate	Yes

Examining the substantive shifts for in-group bias we see that all three Presidents who showed statistically significant variation also show substantive variation. President’s Bush, Johnson and Truman all move from the Low ranking to a higher ranking when moving from spontaneous to prepared material.

Table 3.12d Substantive Shifts for In-group Bias

Trait	Pres	Spontaneous Category	Prepared Category	Change?
IGB	Bush	Low	High	Yes
	Johnson	Low	Moderate	Yes
	Truman	Low	High	Yes

With the need for power trait we see mixed results, Presidents Clinton and Truman show a change from the Moderate rank to the High ranking when moving from spontaneous to prepared material. Johnson, on the other hand, remains Moderate regardless of the material used to produce need for power scores.

Table 3.12e Substantive Shifts for Need for Power

Trait	Pres	Spontaneous Category	Prepared Category	Change?
PWR	Clinton	Moderate	High	Yes
	Johnson	Moderate	Moderate	No
	Truman	Moderate	High	Yes

The remaining two traits, self-confidence and task orientation both show statistically significant variation for all four Presidents. Looking at self-confidence we also see that each President has a negative shift in rank when comparing scores based on spontaneous versus prepared material. The same holds true for the task orientation trait where all four Presidents move to a lower rank when comparing scores based on spontaneous versus prepared material.

Table 3.12f Substantive Shifts for Self-Confidence

Trait	Pres	Spontaneous Category	Prepared Category	Change?
SC	Bush	High	Low	Yes
	Clinton	High	Moderate	Yes
	Johnson	High	Moderate	Yes
	Truman	High	Moderate	Yes

Table 3.12g Substantive Shifts for Task Orientation

Trait	Pres	Spontaneous Category	Prepared Category	Change?
TASK	Bush	Moderate	Low	Yes
	Clinton	Moderate	Low	Yes
	Johnson	Moderate	Low	Yes
	Truman	High	Moderate	Yes

While these substantive differences in trait scores implies that prepared material is not an accurate reflection of a leader’s personality we may still be able to use prepared material. Because the scores for each trait are only meaningful in comparison to other leaders, we may find that the patterns in direction for each trait lead to similar classifications for each President as long as prepared scores are compared to the mean score of the population of leaders using only prepared remarks (and the same with spontaneous material). The data in Tables 3.12a – g is compared to the population mean and standard deviation scores for all data (and not data split by the spontaneity of material). To test whether this influenced the findings, an additional test was conducted examining the trait classifications for each leader based on comparisons of spontaneous leader data to the population of spontaneous material and prepared leader data to the population of prepared material.

AD HOC Hypothesis 1: Trait score classifications derived from a leader’s prepared material compared to population means scores derived from only prepared material will be identical to trait score classifications derived from a leader’s spontaneous material compared to population mean scores derived from only spontaneous material.

As Table 3.13 shows, AD HOC Hypothesis 1 is false. Each of the Presidents included in this analysis are classified differently on a number of traits even when the comparison base to develop classifications is based on the same type of material. Bush would be classified differently on two traits, Clinton for three traits, Johnson for four traits, and Truman on five

traits. Therefore, in the absence of spontaneous material, it is not advisable to use prepared material compared to the population of prepared material as a proxy for spontaneous material.

Table 3.13 Comparison of Trait Classifications based on Different Population Material

Trait	Pres	Spontaneous Data	Prepared Data	Change?
BACE	Bush	Moderate	Moderate	No
	Clinton	Moderate	High	Yes
	Johnson	Moderate	Moderate	No
	Truman	Moderate	Moderate	No
CC	Bush	Moderate	Moderate	No
	Clinton	Moderate	Moderate	No
	Johnson	Moderate	Moderate	No
	Truman	Low	Moderate	Yes
DIS	Bush	Moderate	Moderate	No
	Clinton	Moderate	Moderate	No
	Johnson	Low	Moderate	Yes
	Truman	Low	Moderate	Yes
IGB	Bush	Moderate	Moderate	No
	Clinton	Moderate	Low	Yes
	Johnson	Low	Moderate	Yes
	Truman	Low	High	Yes
PWR	Bush	Moderate	Moderate	No
	Clinton	High	High	No
	Johnson	Moderate	Moderate	No
	Truman	Moderate	Moderate	No
SC	Bush	High	Moderate	Yes
	Clinton	High	Moderate	Yes
	Johnson	High	Moderate	Yes
	Truman	Moderate	High	Yes
TASK	Bush	Moderate	Low	Yes
	Clinton	Moderate	Moderate	No
	Johnson	Moderate	Low	Yes
	Truman	High	Moderate	Yes

While testing each President’s spontaneous and prepared material against the population of spontaneous versus the population’s prepared material is useful for assessing differences in scores according to the accepted method for classification, it does not assess whether the

differences between the two types of material are consistent and thus useful as a proxy of one another. For example, while the test of Hypothesis 1 confirmed that a difference exists between spontaneous and prepared data, it does not assess whether that difference is consistent from one response to the next. If the difference between spontaneous and prepared material always differs by .25 for any given trait, that difference can be used to adjust prepared material and thus increase the total amount of data available for analysis. By aggregating trait scores across time, a test can be performed to assess the level of correlation between the two types of material. The correlation was performed twice for each of the four Presidents, first by year and then by quarter. Table 3.14 displays the results from this analysis.

Table 3.14 Correlations between spontaneous and prepared data by Year and by Quarter.

President	Trait	Correlation by Year	Correlation by Quarter
Bush	BACE	0.00	-0.04
	CC	-0.54	0.00
	DIS	-0.14	-0.05
	IGB	0.90	0.16
	PWR	0.99	0.28
	SC	0.27	0.49
	TASK	0.32	0.57
Clinton	BACE	0.98	-0.14
	CC	0.34	0.15
	DIS	-0.56	0.25
	IGB	-0.68	-0.21
	PWR	0.28	-0.17
	SC	0.03	-0.04
	TASK	0.85	0.13
Johnson	BACE	-0.58	0.15
	CC	0.43	0.11
	DIS	0.97	0.64
	IGB	0.95	0.38
	PWR	-0.95	-0.32
	SC	0.12	0.02
	TASK	0.69	0.04
Truman	BACE	0.53	-0.02
	CC	0.81	-0.14

	DIS	0.30	0.14
	IGB	-0.97	0.25
	PWR	-0.22	-0.13
	SC	-0.58	-0.11
	TASK	-0.65	0.04

Examining the correlation we find that there is considerable correlation between spontaneous and prepared scores aggregated by year. In nine of the twenty-eight cases the level of correlation between the two types of scores exceeds .80. Moving to data aggregated by quarter years, however, leads to a much different result. In no cases does the correlation exceed .80, with the highest level of correlation between Johnson’s distrust of others scores at .64. In only five cases does the score exceed .30. Because of this difference it is safe to assume that a significant portion of the correlation between scores at the year level is a result of collapsing the data into four data points. The low level of correlation found at the quarter level of aggregation lends further support to the claim that the two types of material differ and that Hermann’s suggestion to limit profiles to spontaneous material should be followed.

The last test to be conducted in this chapter is an assessment of difference in trait scores by leader. As is the case with all approaches in the foreign policy decision-making school, finding differences across individuals is essential to the validity of the Leadership Trait Analysis technique. After all, if there is no observable difference from one leader to the next then there is no basis for explaining differences in behavior. Hypothesis 3 will be tested to assess whether President is a significant source of variation in scores across the entire dataset:

Hypothesis 3: President is a significant source of variation in trait scores across all data coded in this study.

Performing ANOVA on all spontaneous responses for the four leaders included in this study, we find that all but one trait shows statistically significant variation by leader (see Table 3.14 for details).

Table 3.15 Results of ANOVA by President on all Spontaneous Data

Trait	DF	Probability
BACE	3	< .0001
CC	3	< .0001
DIS	3	< .0001
IGB	3	.2597
PWR	3	< .0001
SC	3	< .0001
TASK	3	< .0001

From Table 3.14 we can see that in-group bias fails to report any significant difference in scores by President. Before we reject the validity of the in-group bias trait (at least for the four leaders examined in this study), there are two notes of caution. First, this result may be a product of the automated coding process. In order to code in-group bias, references to the in-group of each leader need to be identified so that the system can then assess whether the leader mentions those in-groups in some sort of glorified manner. The Profiler Plus content analysis software contains a function that identifies all potential in-group candidates found within the text coded for a leader. Performing this function produces a list of in-group candidates with multiple thousands of entries. Because the computer has no way to automatically identify who is and is not a member of any particular individual's in-group, we must go through the list and identify all candidates which should be coded as an in-group for the leader in question. Because of the volume of data coded in this analysis, time and labor constraints made the identification of all in-groups unfeasible for this analysis. All references to the United States and the U.S. presidency, however, were identified as in-groups because this remained constant across all Presidents. In addition, the coding scheme looks for all generic references to in-groups (such as "we", "us", or "our") that are universally applicable.

The second factor that may explain the lack of difference in scores may be found in the next chapter. If response topic is found to be a significant source of variation for the in-group bias trait, the lack of difference between leaders observed here may be masked by the relative distribution of response topics within each leader's dataset. In fact, this concern may also have produced a false confirmation of significant variation for the other traits. If response topic does provide a significant source of variation for each of the Presidents in this study, the relative distribution of the response topics within their dataset could make differences appear that would disappear when the data is segmented by response topic. Chapter four will assess this possibility in Hypotheses 6a-c by performing ANOVA tests for President on the data segmented by the three response topics that are of substantive interest to scholars of foreign policy behavior (conflict/security, economics, and politics). This preliminary assessment, however, does provide additional support to the claim that the Leadership Trait Analysis technique is capable of identifying differences in trait scores for different leaders.

Conclusion

As expected, both statistically and substantively significant differences were found between the majority of the prepared and spontaneous trait scores for each of the four Presidents included in this study. While some traits for some leaders did not appear to vary by spontaneity, there was statistically significant variation in trait scores for at least three of the four Presidents on each trait (with self-confidence and task orientation showing statistically significant differences for all four Presidents). For the three Presidents who failed to show significant variation on all seven traits, there is no evidence to support the position that this results from the skillful use of instrumental speech. Using a population of world leaders as a comparison base and

following the technique employed by Hermann to categorize leaders into High, Moderate and Low categories, we find that there is also substantive differences for each personality trait measured with Leadership Trait Analysis. A clear pattern was identified in the direction of the differences between trait scores derived from prepared material versus the trait scores produced from spontaneous material (with the exception of distrust of others). This warrants further research to determine if this is simply a coincidence with the four leaders in this study or if a broader pattern is present. The consistent pattern in the direction of these differences may tell us something about the messages and signals that world leaders attempt to convey in their prepared remarks. The fact that the magnitude of the differences between mean trait scores derived from prepared versus spontaneous materials was not uniform or predictable within any single leader's mean trait scores confirms Hermann's skepticism regarding the use of non-spontaneous material in the production of trait scores.

Because significant variation was found between trait scores based on prepared versus spontaneous material, the remainder of this analysis will be limited to only spontaneous material. While some traits for some leaders did not show significant variation based on spontaneity, the lack of consistency in differences for specific traits across leaders leaves open the possibility that the similarities were. Without detailed data regarding the author(s) of prepared speeches and the degree of involvement by the President in its creation, we are left with continued skepticism regarding the use of prepared material in the production of personality trait scores using the Leadership Trait Analysis technique.

Chapter 4

State versus Trait I Trait Score Difference by Response Topic

Both the Leadership Trait Analysis and foreign policy decision-making literature as a whole have paid considerable attention to issues of instrumental speech. Within the Leadership Trait Analysis research program, Brian Dille and Michael D. Young have examined whether the level of spontaneity of utterances by world leaders is a source of significant variation for personality trait scores. While these tests were not conclusive, they did take the first steps toward critically reviewing the data used in personality assessment.

While a number of substantively important sources of potential variation exists (such as testing for variance based on the location where the speech was delivered, who was in the audience, what type of news media covered the event, etc...) a potentially significant source of variation has yet to be tested: topical content of remarks. We might assume that a leader talking about war against an evil dictator would appear more distrustful than when that same leader is discussing the government's role in the economy. On the other hand, it might be argued that whether you are dealing with politics, security, or the economy, there are ample opportunities within each issue to use words which would make a leader appear more or less conceptually complex, distrustful of others, or have a higher or lower need for power. To date, there have been no published tests that examine whether the topic of the material used with the Leadership Trait Analysis technique is a source of significant variation in trait scores. This chapter will take a first look at this issue.

Preparation of the press conferences and interviews included the assignment of a response topic for each answer to questions fielded by the Presidents in this study. Four

categories were used to classify each response: conflict/security, economics, politics, and “other”. Conflict/security topics include any discussion of armed conflicts, negotiations for settling such conflicts, defense policy, and civil unrest. Economic topics include discussions of foreign trade, international financial institutions, and the domestic economy. Political topics include any discussion of government policy, initiatives, or legislation that do not fit into the preceding categories. The “other” category includes any discussion of personal issues, logistical statements during the interview or press conference (such as questions of clarification or announcements regarding the number of questions left before the President will depart, etc...) or other topics that do not fall into the other three categories. Responses to questions consisted of multiple topics were split into multiple responses only at the paragraph breaks located in the text provided by the Public Papers of the Presidents series. If multiple topics were identified within a single paragraph, a coding hierarchy with conflict/security issues at the top, economic issues second, and political issues third was used to make a coding decision. While records were not kept detailing the coding process, I would estimate that at least 80% of the topics fit clearly into one category, with the remaining 20% capable of being classified in at least two of the categories.

Given the findings of Chapter 3, this study is operating under the assumption that spontaneous material better represents a leader's personality. Therefore the remaining hypotheses in this analysis will be tested using only spontaneous material¹⁰.

Hypothesis 4: Mean trait scores for each President will exhibit statistically significant variation across response topic.

¹⁰ Limiting the sample for each leader to only spontaneous material reduces the sample size yet still offers significantly more data than has traditionally been used in studies employing the Leadership Trait Analysis technique.

If statistically significant variation in mean trait scores is found across response topic, Hypothesis 5 will then be tested.

Hypothesis 5: Trait scores which exhibit statistically significant variation based on response topic will also exhibit substantively significant variation.

If both statistically significant and substantive variation is found based on response topic, then trait stability over time will be tested separately for each topic in Chapter 5.

Findings

Table 4.1 displays the results of the ANOVA test for response topic for George Bush, Sr. Of the seven personality traits measured by Leadership Trait Analysis, six varied significantly based on response topic. In-group bias just exceeds the .05 level of significance threshold with a probability of .053. Of the significant categories, belief in one's ability to control events shows the greatest amount of variation between topics with a .12 difference in trait scores between economic responses (at the high end with a mean score of .59) and "other" responses (at the low end with a mean score of .47). Both need for power and belief in one's ability to control events show the least amount of variation between topics with a .04 difference in mean trait scores between high and low categories.

Table 4.1 Response Topic ANOVA for George Bush, Sr.

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	< .01	Conflict / Security	1789	176,701	0.30
		Economics	899	125,847	0.34
		Politics	2933	340,227	0.31
		Other	481	46,442	0.30
CC	< .001	Conflict / Security	1922	197,053	0.61
		Economics	926	132,509	0.66
		Politics	3058	370,689	0.64
		Other	604	49,479	0.63
DIS	< .005	Conflict / Security	1502	119,574	0.11
		Economics	782	90,577	0.09
		Politics	2518	228,120	0.10
		Other	409	39,533	0.15
<i>IGB</i>	<i>.053</i>	Conflict / Security	1261	98,442	0.06
		Economics	689	82,904	0.07
		Politics	1984	194,427	0.07
		Other	235	34,018	0.09
PWR	< .001	Conflict / Security	1787	168,318	0.24
		Economics	893	124,256	0.28
		Politics	2912	328,508	0.25
		Other	477	46,740	0.27
SC	< .001	Conflict / Security	1732	174,852	0.56
		Economics	868	123,209	0.59
		Politics	2873	337,815	0.54
		Other	488	48,018	0.47
TASK	< .001	Conflict / Security	1691	176,477	0.58
		Economics	868	125,183	0.66
		Politics	2780	337,155	0.61
		Other	451	45,342	0.57

Table 4.2 displays the results of the ANOVA test by response topic for Bill Clinton. Like Bush, six of the seven traits for Clinton show statistically significant variation by response topic. Belief in one’s ability to control events fails to show significant variation with a probability of .5. Of the significant categories, self-confidence shows the greatest amount of variation between topics with a .24 difference in trait scores between conflict/security responses (at the high end with a mean score of .63) and “other” responses (at the low end with a mean score of .39). Both distrust of others and in-group bias possess the least amount of variation between topics with a .04 difference in mean trait scores between high and low categories.

Table 4.2 Response Topic ANOVA for Bill Clinton

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
<i>BACE</i>	.5	Conflict / Security	1623	154,056	0.36
		Economics	812	103,346	0.35
		Politics	1135	164,214	0.36
		Other	393	27,975	0.34
CC	< .001	Conflict / Security	1623	161,610	0.58
		Economics	833	107,880	0.65
		Politics	1147	167,971	0.57
		Other	405	28,332	0.50
DIS	< .01	Conflict / Security	1355	95,142	0.08
		Economics	729	69,077	0.09
		Politics	992	131,097	0.10
		Other	309	23,408	0.12
IGB	< .05	Conflict / Security	1385	87,179	0.05
		Economics	737	66,670	0.05
		Politics	854	117,389	0.07
		Other	230	20,358	0.09

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
PWR	< .005	Conflict / Security	1613	147,929	0.26
		Economics	810	98,226	0.26
		Politics	1124	158,736	0.26
		Other	383	27,212	0.21
SC	< .001	Conflict / Security	1491	140,601	0.63
		Economics	753	93,982	0.60
		Politics	1101	159,093	0.56
		Other	397	27,289	0.39
TASK	< .01	Conflict / Security	1576	155,697	0.63
		Economics	812	102,834	0.66
		Politics	1083	160,310	0.62
		Other	296	25,958	0.61

Table 4.3 shows the results of the ANOVA test by response topic for Lyndon Johnson. Unlike Presidents Bush and Clinton, only three of Johnson's seven traits exhibit statistically significant variation by response topic. Belief in one's ability to control events, in-group bias, need for power, and self-confidence fail to show significant variation. Of the three significant categories, task orientation shows the greatest amount of variation between topics with a .12 difference in trait scores between economic responses (at the high end with a mean score of .70) and other responses (at the low end with a mean score of .58). Distrust of others shows the least amount of variation between topics with a .07 difference in mean trait scores between high and low categories.

Table 4.3 Response Topic ANOVA for Lyndon Johnson

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
<i>BACE</i>	.06	Conflict / Security	485	75,011	0.34
		Economics	359	63,672	0.38
		Politics	732	96,238	0.33
		Other	85	11,239	0.38
CC	< .001	Conflict / Security	468	76,431	0.56
		Economics	350	63,047	0.65
		Politics	675	98,398	0.56
		Other	77	11,179	0.58
DIS	< .001	Conflict / Security	320	61,964	0.10
		Economics	246	52,570	0.06
		Politics	526	71,976	0.04
		Other	47	8,818	0.03
<i>IGB</i>	.1	Conflict / Security	369	64,652	0.07
		Economics	297	57,083	0.04
		Politics	424	72,905	0.06
		Other	44	8,790	0.08
<i>PWR</i>	.48	Conflict / Security	482	73,726	0.22
		Economics	358	62,750	0.24
		Politics	726	93,224	0.22
		Other	84	11,164	0.20
SC	.06	Conflict / Security	444	71,611	0.58
		Economics	304	57,667	0.57
		Politics	692	93,590	0.55
		Other	84	11,145	0.48
TASK	< .01	Conflict / Security	432	73,442	0.64
		Economics	320	61,203	0.70
		Politics	607	92,177	0.63
		Other	66	10,691	0.58

Table 4.4 displays the results of the ANOVA test by response topic for Harry Truman. Truman shows significant variation by response topic for four of the seven traits measured by Leadership Trait Analysis. Conceptual complexity, in-group bias, and distrust of others fail to show significant variation by response topic. Of the four significant categories, task orientation shows the greatest amount of variation between topics with a .22 difference in trait scores between economic and political responses (both at the high end with a mean score of .77) and other responses (at the low end with a mean score of .55). Need for power shows the least amount of variation between topics with a .06 difference in mean trait scores between economic responses (at the high end with a mean score of .27) and conflict/security responses (at the low end with a mean score of .21).

Table 4.4 Response Topic ANOVA for Harry Truman

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
BACE	< .005	Conflict / Security	422	10,123	0.33
		Economics	815	16,825	0.38
		Politics	1639	22,189	0.36
		Other	528	9,610	0.31
CC	.11	Conflict / Security	359	11,068	0.45
		Economics	860	18,666	0.50
		Politics	1465	24,210	0.47
		Other	552	10,349	0.49
DIS	.14	Conflict / Security	268	6,762	0.02
		Economics	490	8,878	0.04
		Politics	1085	13,579	0.04
		Other	333	5,282	0.05

Trait	Probability	Spontaneity	Response Count	Word Count	Mean
IGB	.94	Conflict / Security	122	4,930	0.06
		Economics	263	8,527	0.05
		Politics	282	9,168	0.06
		Other	66	4,104	0.05
PWR	< .005	Conflict / Security	389	9,102	0.21
		Economics	767	14,767	0.27
		Politics	1549	20,326	0.22
		Other	505	8,252	0.22
SC	< .001	Conflict / Security	402	9,536	0.57
		Economics	770	15,354	0.56
		Politics	1724	22,732	0.52
		Other	547	9,749	0.48
TASK	< .001	Conflict / Security	282	8,221	0.74
		Economics	570	13,051	0.77
		Politics	1022	18,295	0.77
		Other	345	6,988	0.55

Analysis

The data found in Tables 4.1 through 4.4 clearly show that response topic is a significant source of the variation found in trait scores as measured by the Leadership Trait Analysis technique. While we failed to find significant variation for all traits for all Presidents, 68% of the traits did show significant variation. Of the traits that did not have significant variation, there does not appear to be any discernible pattern as to why each trait for each leader was or was not significant – with the possible exception of in-group bias that exhibited insignificant variation for three of the four Presidents (see Table 4.5 for a summary of significant variation). Among Presidents, we see that Bush and Clinton both show variation in six of the seven traits, with in-group bias being insignificant for Bush and belief in one’s ability to control events for Clinton.

Of Johnson’s seven personality traits only three traits showed significant variation based on response topic with belief in one’s ability to control events, in-group bias, need for power, and self-confidence all failing to show statistically significant variation by response topic. Truman showed slightly more variation with four of the seven traits being significant (conceptual complexity, distrust of others, and in-group bias failed to show statistically significant variation based on response topic). Taken as a whole, this leaves us with only one trait where three Presidents failed to show significant variation (in-group bias), and one trait where two Presidents failed to show significant variation (belief in one’s ability to control events). For each of the remaining traits only one President failed to show significant variation by response topic. While further analysis of additional U.S. Presidents and other world leaders is suggested, these initial findings support Hypothesis 4.

Table 4.5 Summary of Statistically Significant Variation on Response Topic

Trait	Bush	Clinton	Johnson	Truman
BACE	Yes	No	No	Yes
CC	Yes	Yes	Yes	No
DIS	Yes	Yes	Yes	No
IGB	No	Yes	No	No
PWR	Yes	Yes	No	Yes
SC	Yes	Yes	No	Yes
TASK	Yes	Yes	Yes	Yes

As was done in the last analysis of trait score variation by spontaneity, we must also test for the existence of substantive variation by response topic. While the last chapter performed the substantive comparison to all data from the population of leaders provided by Social Science Automation, the comparison in this chapter will only be compared to population data calculated

by coding spontaneous material. The population mean scores and standard deviations for each trait based only on spontaneous material is reported in Table 4.6 below.

Table 4.6 Population Mean Trait Scores and Comparison Ranges*

Trait	Population Mean	Standard Deviation	Low Range	Moderate Range	High Range
BACE	.34	.05	<= .29	>.29 and <.39	>= .39
CC	.59	.05	<= .54	>.54 and <.64	>= .64
DIS	.12	.04	<= .08	>.08 and <.16	>= .16
IGB	.08	.02	<= .06	>.06 and <.10	>= .10
PWR	.23	.03	<= .20	>.20 and <.26	>= .26
SC	.46	.09	<= .37	>.37 and <.55	>= .55
TASK	.65	.06	<= .59	>.59 and <.71	>= .71

* data in this table derived from spontaneous material only

While four categories were used to determine whether statistically significant variation existed (conflict/security, economics, politics, and other) the test for substantive variation will exclude the “other” category because it is of little substantive interest. The “other” category is a collection of issues discussed by leaders that do not fall into the other three categories. Because the foreign policy decision-making approach is primarily interested in applying findings to policy situations, this analysis is limited to the areas of primary concern in foreign affairs: economic, political, and security relations among states¹¹.

¹¹ The test of statistical significance was performed a second time excluding the other category in the analysis. The findings were nearly identical with only four additional leader-traits becoming insignificant for 54% of all traits showing significant variation based on response topic.

Using the data in Table 4.6 we can classify each of Bush’s trait scores for each category into High, Moderate, or Low categories for comparison. Table 4.7 shows the results of the substantive test of variation for George Bush, Sr. based on response topic. We see substantive variation in four of the six traits that exhibited statistically significant variation above. Both belief in one’s ability to control events and distrust of others fails to show substantive variation across response topic.

Table 4.7 Substantive Variation by Response Topic for George Bush, Sr.*

Trait	Conflict/Security	Economics	Politics	Change?
<i>BACE</i>	<i>Moderate</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
CC	Moderate	High	High	Yes
<i>DIS</i>	<i>Moderate</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
PWR	Moderate	High	Moderate	Yes
SC	High	High	Moderate	Yes
TASK	Low	Moderate	Moderate	Yes

* only traits which showed statistically significant variation are shown

Table 4.8 displays the results of the substantive test of variation for Bill Clinton based on response topic. We see substantive variation in three of the six traits that exhibited statistically significant variation above. Need for power, self-confidence and task orientation all fail to show substantive variation across response topic.

Table 4.8 Substantive Variation by Response Topic for Bill Clinton*

Trait	Conflict/Security	Economics	Politics	Change?
BACE	Moderate	High	Moderate	Yes
CC	Low	Moderate	Moderate	Yes
DIS	Low	Low	Moderate	Yes
<i>PWR</i>	<i>High</i>	<i>High</i>	<i>High</i>	<i>No</i>
SC	<i>High</i>	<i>High</i>	<i>High</i>	<i>No</i>
TASK	<i>Moderate</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>

* only traits which showed statistically significant variation are shown

Table 4.9 displays the results of the substantive test of variation for Lyndon Johnson based on response topic. Johnson only showed significant variation in three trait scores, leaving us with only conceptual complexity, distrust of others and task orientation. Of these three traits, conceptual complexity and distrust of others showed substantive variation. Task orientation, on the other hand, failed to show substantive variation between response topic categories.

Table 4.9 Substantive Variation by Response Topic for Lyndon Johnson *

Trait	Conflict/Security	Economics	Politics	Change?
CC	Moderate	High	Moderate	Yes
DIS	Moderate	Low	Low	Yes
<i>TASK</i>	<i>Moderate</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>

* only traits which showed statistically significant variation are shown

Table 4.10 displays the results of the substantive test of variation for Harry Truman based on response topic. Truman only showed significant variation in four trait scores, leaving us with belief in one's ability to control events, need for power, self-confidence and task orientation. Of these four traits, need for power and self-confidence showed substantive variation. Belief in one's ability to control events and task orientation, on the other hand, failed to show substantive variation between response topic categories.

Table 4.10 Substantive Variation by Response Topic for Harry Truman *

Trait	Conflict/Security	Economics	Politics	Change?
<i>BACE</i>	<i>Moderate</i>	<i>Moderate</i>	<i>Moderate</i>	<i>No</i>
PWR	Moderate	High	Moderate	Yes
SC	High	High	Moderate	Yes
<i>TASK</i>	<i>High</i>	<i>High</i>	<i>High</i>	<i>No</i>

* only traits which showed statistically significant variation are shown

Of the nineteen traits reporting statistically significant variation for the four Presidents, eleven also showed substantive variation between the three topics of interest deemed relevant to the study of foreign policy decision-making (conflict/security, economics, and politics). Closer examination of the substantive variation observed shows some pattern within traits, however it is questionable as to whether a pattern exists within President. For conceptual complexity (which showed statistically significant variation for three Presidents), all of the Presidents showed substantive variation, with economic responses scoring higher than at least one other category and never less than another category. Task orientation (which showed statistically significant variation for all four Presidents) showed substantive variation for only one President. Belief in one's ability to control events (where only two Presidents showed significant statistical variation) failed to show substantive variation in both cases. Distrust of others, need for power, and self-confidence each had three Presidents who showed statistical variation and in each case two Presidents showed substantive variation as well.

Within President the only pattern that is observed is that economic responses tend to be classified in a higher category than the political or conflict/security responses. Both conflict/security and political responses' relative comparison to the other two categories appear to fluctuate randomly. While we see substantive variation for nine of the twenty-eight leader-trait combinations measured in this study, we are left with little ability to predict when that variation will occur for any leader and for any specific trait. These results are at the very least a partial confirmation of Hypothesis 5, and the next chapter's analysis of temporal stability should be conducted separately for each response topic.

The final hypotheses to be tested in this chapter are a follow-up to the final hypothesis of Chapter 3, examining whether President serves as a significant source of variation among trait

scores. Because significant differences were found by response topic, this test will be performed separately for each of the three response topics of substantive interest. Hypotheses 6a-c are listed below:

Hypothesis 6a: President is a significant source of variation in trait scores across all conflict/security responses coded in this study.

Hypothesis 6b: President is a significant source of variation in trait scores across all economic responses coded in this study.

Hypothesis 6c: President is a significant source of variation in trait scores across all political responses coded in this study.

Performing ANOVA on all responses within each topic we find that the majority of traits show significant variation by leader for the three response topics assessed in this study. The only trait that did not show statistically significant variation by President for all three topics was in-group bias, which adds further support to the argument that in-group bias is not measuring anything distinct across individuals. Further research, however, must be conducted to assess whether in-group bias' failure to show statistically significant differences in this analysis is a product of a) the small sample of leaders included in this study or b) neglecting to perform in-group assignments for the in-group candidates identified by Profiler Plus.

For conflict and political responses, all of the remaining traits show statistically significant variation by President. For economic responses we fail to see statistically significant difference on two additional traits. In addition to in-group bias, need for power and self-confidence fail to show significant variation ($p < .05$). This leaves us with sixteen of twenty-one trait-topic combinations showing statistically significant variation by President.

Table 4.11 ANOVA Probability findings for variation in trait scores by President for each topic

Topic	Trait	Probability
Conflict	BACE	< .0001
	CC	< .0001
	DIS	< .0001
	<i>IGB</i>	<i>0.67</i>
	PWR	< .001
	SC	< .0001
	TASK	< .0001
Economic	BACE	< .05
	CC	< .0001
	DIS	< .0001
	<i>IGB</i>	<i>.14</i>
	<i>PWR</i>	<i>.15</i>
	<i>SC</i>	<i>.11</i>
	TASK	< .0001
Political	BACE	< .0001
	CC	< .0001
	DIS	< .0001
	<i>IGB</i>	<i>.60</i>
	PWR	< .0005
	SC	< .05
	TASK	< .0001

The failure to show significant difference in scores by President for self-confidence and need for power on economic topics would also benefit from future research to assess whether this is a function of the small sample of Presidents included in this study or that leaders show less variation on those traits in economic situations.

Conclusion

Both statistically and substantively significant variation was found for a number of traits for each of the four Presidents included in this analysis. These results lead to the conclusion that response topic is a significant source of variation and all of the following tests in this study will divide the data for each President by response topic. Unfortunately, the variation (either statistical or substantive) was not found for all traits for all Presidents. In addition, no clear

pattern can be identified to predict when a trait for a given leader will not show significant variation (either statistical or substantive).

Further research is necessary to assess whether this absence of a pattern is a product of the small sample of total leaders used in this analysis (while the total data used to calculate scores far outweighs any studies done in the past, the total number of leaders studied is still only four). Performing this same analysis on the remaining post World War II U.S. Presidents or a larger sample of world leaders may provide additional insight on this matter.

Convincing evidence was not found to support the claim that variation between topics is consistent across traits or across leaders. With the exception of economic responses scoring higher or equal to other categories in most of the significant, political and conflict/security responses failed to consistently place higher or lower in comparison to one another. This too may benefit from a larger sample of leaders to determine if a broader pattern exists.

Chapter 5

State versus Trait II: Trait Score Stability Across Time

As the title of this study suggests, a debate exists over the stability of personality traits measured by Leadership Trait Analysis over time. Are the traits measured by Leadership Trait Analysis stable, and thus accurately named “trait”? Or are they simply a reflection of the environment in which a leader finds themselves and thus more aptly referred to as a state of being or “states” as is commonly referred to in the literature. This question lies at the heart of the Leadership Trait Analysis technique. After all, if personality scores vary randomly from one time period to the next, can they really be viewed as a valid measure of an individual’s personality? Lacking a tangible object called personality that can be extracted from a person and measured in a laboratory, we are left with techniques such as Leadership Trait Analysis to try and assess an individual’s personality. The question of trait stability over time really asks whether or not personality exists or whether people are simply a reflection of their environment. If a score changes dramatically from one period to the next, it is unlikely that the technique only measures what is commonly thought of as personality, and is not also reflective of other factors external to the leader.

Recent work by Hermann (1999) has departed from the original conception of traits as being relatively stable throughout a leader’s time in office and opens the possibility of scores shifting in response to a leader’s environment over time. Factors such as new information, shifting roles and learning all have the potential to influence a leader’s approach to the world. As with the other questions addressed by this analysis, previous research on this issue was plagued by limited amounts of data. While no rationale was given for the unit of aggregation, presumably

this was at least one factor leading to assessments of stability on a yearly basis. While aggregating scores by year has the benefit of ensuring each period consists of substantial amounts of data, it also has the effect of reducing all variation in one year to a single data point.

This leaves us with a fundamental question? How should stability be measured? The strictest test of stability would assess whether differences exist between trait scores from response to response. At the other extreme, data for a leader could be divided into two equal time periods and the mean scores from each period could be compared. Between these extremes, we have days, weeks, months, quarters, or years. Choosing between these different time periods for aggregation is an arbitrary process. The larger the time period used to divide data, the greater the likelihood that stability will be found.

Left without a rationale for choosing between these options, this chapter will examine trait stability over time using three distinct units of time – day, quarter, and year. Days are used as the smallest extreme because it is the smallest unit across time by which we have data. Years are used at the other extreme in accordance with previous research while quarter years were chosen as the middle ground through a relatively arbitrary decision. Stability will be tested according to two techniques drawn from the debate between Hermann and Rasler, Thompson and Chester (International Studies Quarterly, 1980). The first technique assesses whether a leader's score shifts significantly across each time period and leads to Hypothesis 7, which examines whether trait scores remain in the same category (High / Moderate / Low) across all time periods.

Hypothesis 7: Mean trait scores for each time period will be classified in the same category ranking¹².

¹² The dataset of world leader trait scores was provided by Social Science Automation for this study.

Alternatively, stability can be viewed relative to other leaders in the world. While scores may shift within a certain range, the scores of leader A on conceptual complexity may always be higher than leader B's conceptual complexity scores. The second hypothesis is drawn from this perspective.

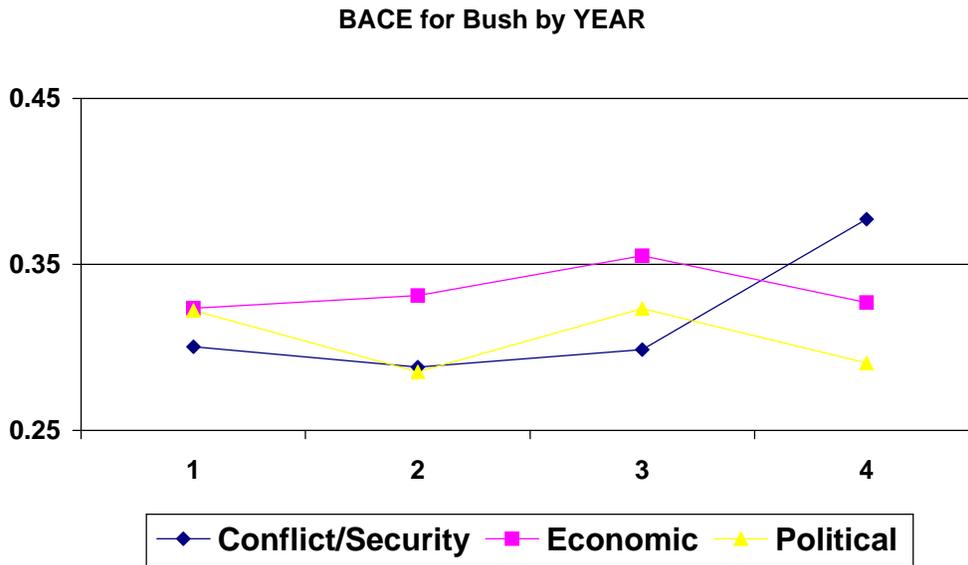
Hypothesis 8: A President's rank for any given trait (compared to the other three Presidents in the study) will remain constant across the entire four years.

The findings for each level of aggregation will be assessed and this chapter will conclude with a discussion of the implications of those results for the future of the Leadership Trait Analysis technique.

Findings

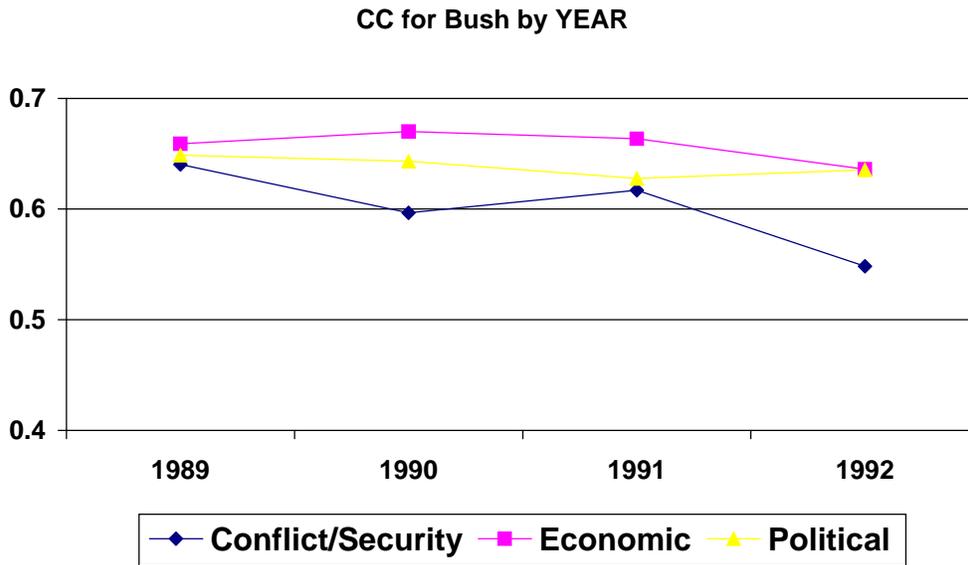
As expected, the aggregation of trait scores by year results in relatively stable trait scores (sharp spikes across the potential range of values are absent) from one period to the next. Using the rankings assigned through comparison of trait scores to the population's belief in one's own ability to control events mean scores and standard deviations, we see mixed results for each leader on each topic. Figure 5.1 shows Bush's mean BACE scores for the three topics used in this analysis aggregated by year. While the scores for each topic remain within a relatively narrow band of all possible values, the mean BACE scores for the Conflict/Security responses does cover the entire span of moderate values (.39 is high, .29 is low). The mean BACE scores for economic responses show less variation across the moderate range but travels close to the high category in the third year. Mean BACE scores for political responses fall solidly within the moderate category across all years.

Figure 5.1



When looking at conceptual complexity (see Figure 5.2), we see less variation from year to year for economic and political responses, yet both topics contain mean scores that are classified in both high and moderate categories (.64 is the boundary between high and moderate classifications). Conflict/Security mean scores for conceptual complexity show more variation from one year to the next and also fall into multiple rank categories with one high classification and three moderate classifications.

Figure 5.2



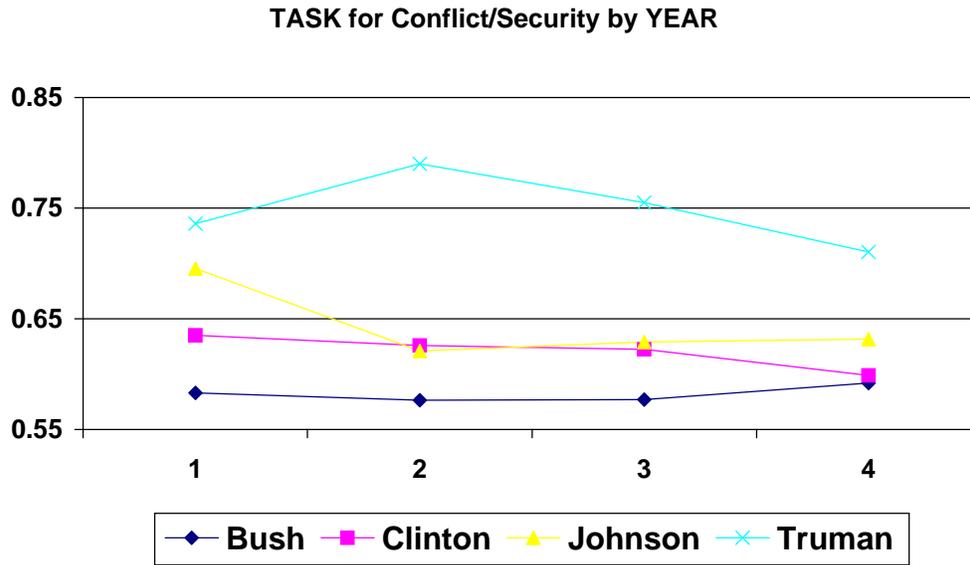
The remaining traits for each President produce similar results. For Bush there are five traits for conflict responses shifting between classifications, three traits for economic responses shifting between classifications and seven traits for political responses shifting between classifications (see Appendix A for tables containing all data for each President). With Clinton's remaining traits we observe five traits for conflict responses shifting between classifications, and four traits each for both economic and political responses shifting between classifications. For Johnson we find five traits for conflict responses shifting between classifications, six traits for economic responses shifting between classifications and four traits for political responses shifting between classifications. With Truman we find four traits each for conflict and economic responses

shifting between classifications and two traits for political responses shifting between classifications.

Looking across all four Presidents we find that at least four of the seven traits derived from conflict responses fail to confirm Hypothesis 7 by displaying shifts in classification from one year to the next. For economic responses we find as few as three traits and as many as six traits fail to confirm Hypothesis 7 and for political responses we find as little as two traits and as many as all seven traits fail to confirm Hypothesis 7. Of the 21 trait-topic combinations for each of the four Presidents (84 total), only twenty-seven remain in the same rank classification across all four periods.

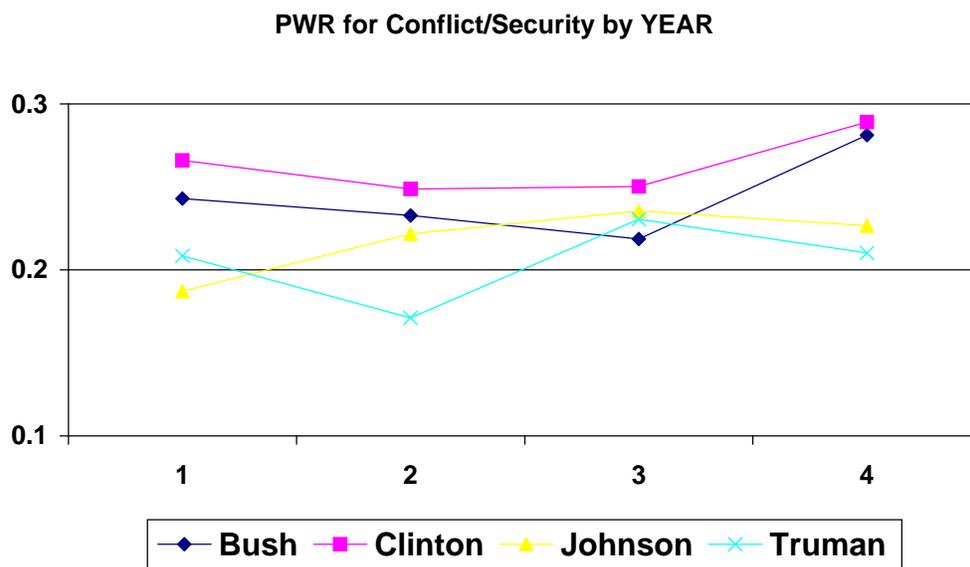
Hypothesis 8, which proposes that the each President's rank relative to each other on a trait will remain constant from one period to the next fails for all traits at the yearly level of aggregation for all topics. The results of this test of Hypothesis 8 can be categorized into three categories. First, four of the twenty-one trait-topic combinations have stability for two leaders (one remains higher than the rest while the other remains lower than the rest). Figure 5.3 shows an example of this category, task orientation for conflict/security responses. In this case Truman is consistently higher than the other three leaders while Bush is consistently lower than the others.

Figure 5.3



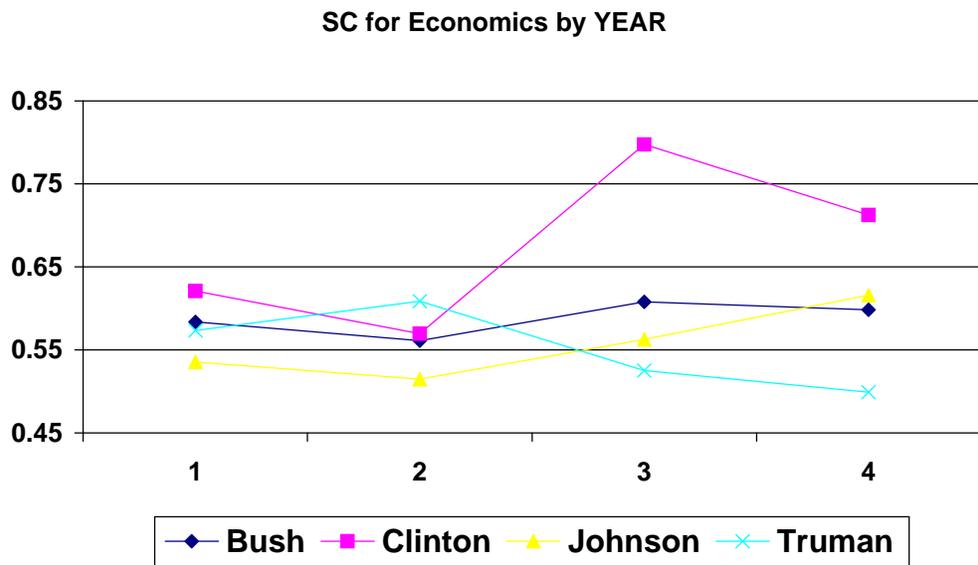
The second category consists of trait-topic combinations where a single leader is consistently higher or lower than the other three Presidents. Of the 21 trait-topic combinations, six fall into this category. Figure 5.4, need for power for conflict responses, shows Clinton scoring higher than the other three Presidents across all time periods.

Figure 5.4



The remaining category consists of the eleven trait-topic combinations that fail to find rank consistency for any of the four Presidents included in this analysis. Figure 5.5, self-confidence for economic responses displays each of the Presidents occupying at least two ranks across the four periods.

Figure 5.5

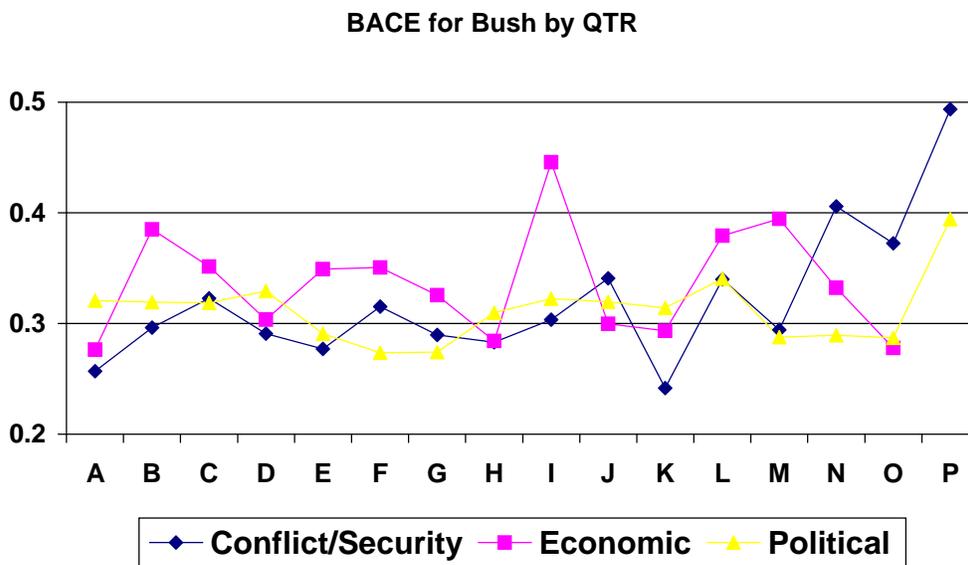


These figures, and the data in Appendix A, show that both Hypotheses 7 and 8 fail to be confirmed. That this result was obtained at the year level of aggregation is quite telling given that aggregation by year greatly minimizes the variation in trait scores observed across the entire period. It is even more of a critical test given that these tests were conducted on reduced comparison sets using only spontaneous material of only one topic further homogenizing the data used to compare scores over time.

As was suspected, the aggregation of trait scores by quarter does result in increased instability in trait scores from one period to the next. Given that there was very little support for

Hypotheses 7 and 8 at the year level, it is not surprising that there is even less support at the quarter level of aggregation. Figure 5.6 shows Bush's mean trait scores on the belief in one's ability to control events trait for the three topics used in this analysis aggregated by quarter. While the same trait showed little spiking from one period to the next at the year level of analysis, considerable movement in mean scores exist for conflict and economic responses when the data is aggregated by quarter. Political responses appear less susceptible to change over time, although there is a distinct rise in mean scores toward the end of the four years. While this initial visual check for stability leads us to believe that Hypothesis 7 would fail for Bush on belief in one's ability to control events, it is still necessary to determine whether the comparative classifications shift across quarter.

Figure 5.6



Checking the mean scores for each trait, Bush has no traits that show consistency of classification across all time periods for any of the three topics. However, in two cases on

economic topics, Bush's trait classification remains the same across the majority of periods. Figures 5.7 and 5.8 show these two traits.

Figure 5.7

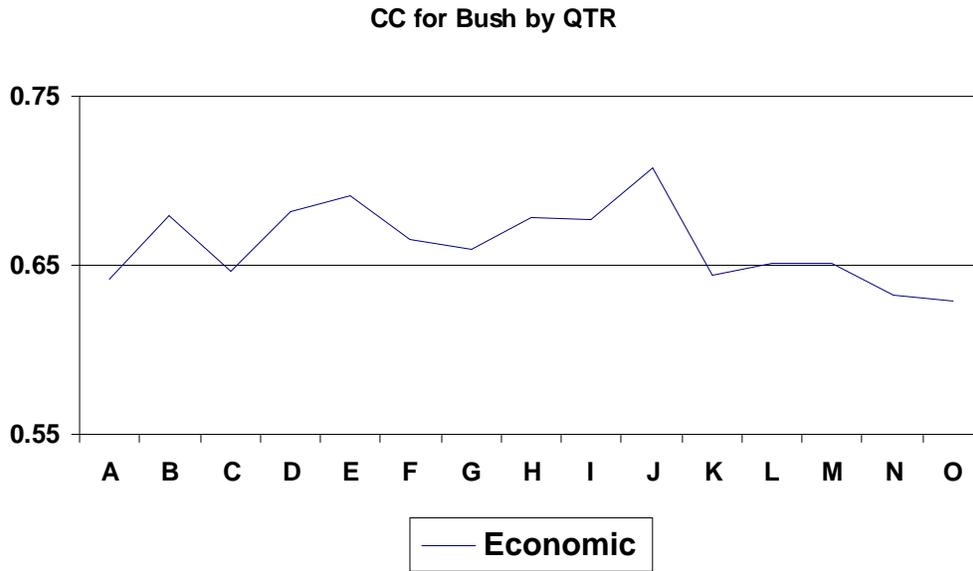
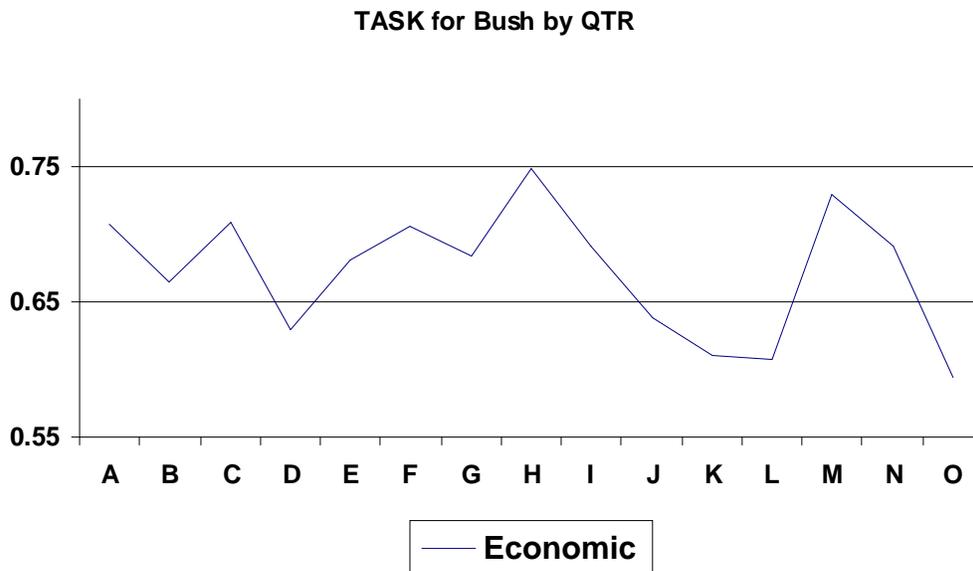


Figure 5.8



In Figure 5.7 we can see that Bush remains in the High category for conceptual complexity for the first three and half years of his administration and it isn't until the last six months that he drops into the moderate category. In Figure 5.8 we see two quarters (H and M) where Bush moves from the moderate task orientation classification into the high classification.

Moving to Clinton we see similar results with no traits remaining in the same classification across all periods. On conflict/security responses, Clinton has two traits where nearly all of the quarters are classified identically, self-confidence and task orientation. On economic responses, Clinton remains in the high classification in fourteen of sixteen periods for self-confidence. There are no traits where Clinton comes close to remaining in the same classification for all periods on political topics.

The data for President Johnson not only fails to confirm Hypothesis 7 for any traits on any topic, but also has no traits that remain in the same classification in at least fourteen of the sixteen quarters, as was witnessed with Bush and Clinton. The data for Truman also fails to confirm Hypothesis 7 but similar to Bush and Clinton, he comes close on a number of traits on different topics. For conflict/security responses, he shows consistency on fourteen of sixteen periods for conceptual complexity and fifteen of sixteen periods of distrust. For economic responses Truman shows consistency on fifteen of sixteen periods for distrust of others. On political responses Truman shows consistency on fifteen of sixteen periods for two traits, conceptual complexity and distrust of others.

Testing Hypothesis 8 for each President, we find that only one of the twenty-one topic-trait combinations has a single President maintaining the same rank across all periods. Truman ranks lower on distrust of others than the other three Presidents on all of the fifteen periods where there is data for each President. Comparisons by quarter for Hypothesis 8 were only tested

on periods where data was available for all four Presidents. Because of the limited number of press conferences during certain quarters for certain Presidents, we are left with a few quarters where there is no score for some topics.

On conflict/security issues, Truman is the only President who maintains a single rank across more than nine periods for any of the traits, besides ranking fourth on distrust for all periods, Truman also ranks fourth for eleven of fifteen periods. On economic responses, Truman again is the only President who maintains a single rank across more than nine periods with a conceptual complexity rank of fourth in eleven of fourteen periods. On political responses, Bush ranks highest in twelve of fifteen periods on conceptual complexity while Truman ranks lowest in twelve of fifteen periods. Truman also ranks highest on political responses for task orientation for thirteen of fifteen periods.

While aggregation by year showed confirmation of Hypothesis 7 in only a few cases, aggregation by quarter further confirms that Hypothesis 7 is disproved. Testing Hypothesis 8 for yearly aggregation, the results showed that nearly half of all trait-topic combinations found at least one President maintaining rank across all four years. Testing Hypothesis 8 on quarterly aggregation units, we see that only a single topic-trait combination had at least one President maintaining rank across all categories. These results clearly disprove Hypothesis 8 as well.

Left with a day to day aggregation of results to test Hypothesis 7 and 9, we see a clear confirmation that Hypotheses 7 and 8 are disproved. Figures 5.9 through 5.12 show a single trait-topic combination (belief in one's ability to control events on conflict/security responses) for each President as an illustration of the clear lack of stability from day to day.

Figure 5.9

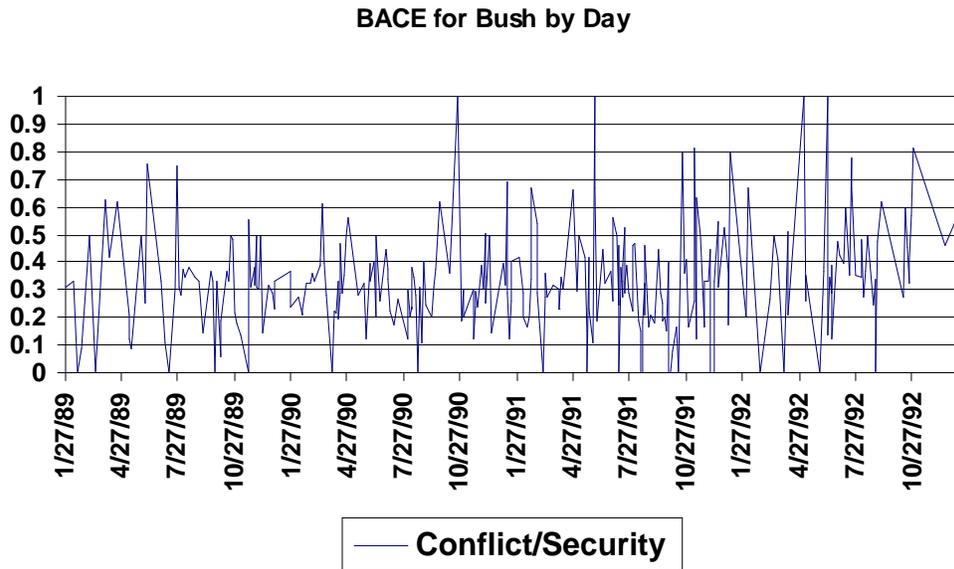


Figure 5.10

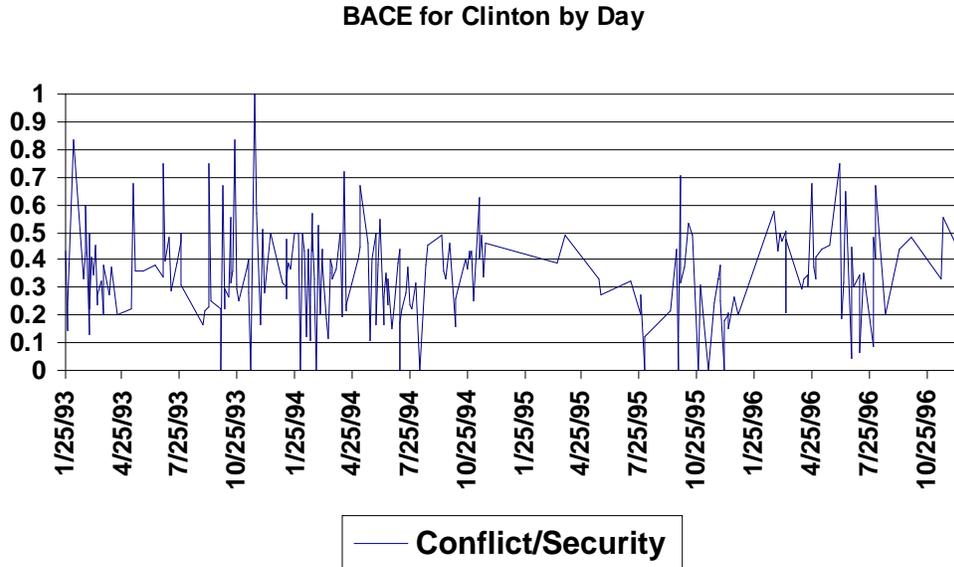


Figure 5.11

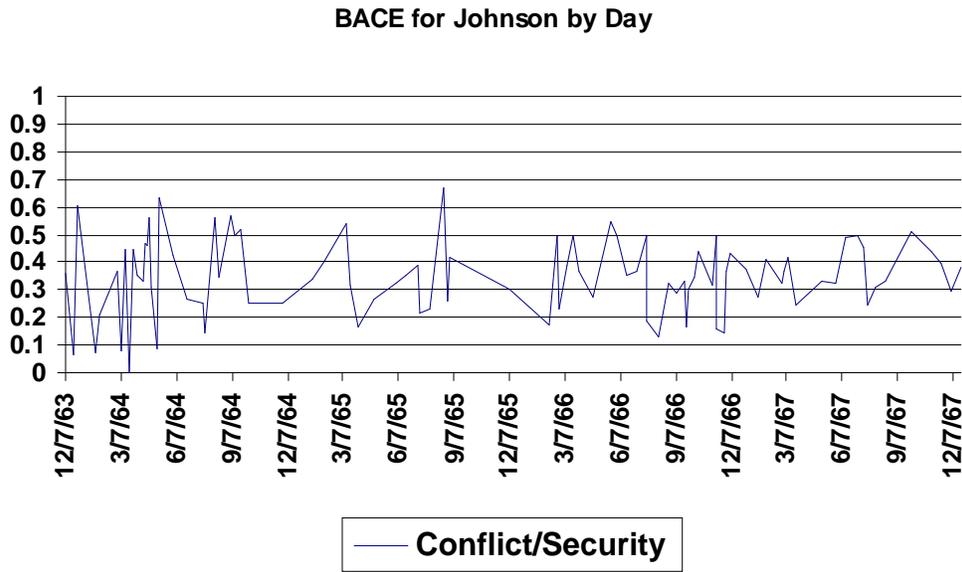
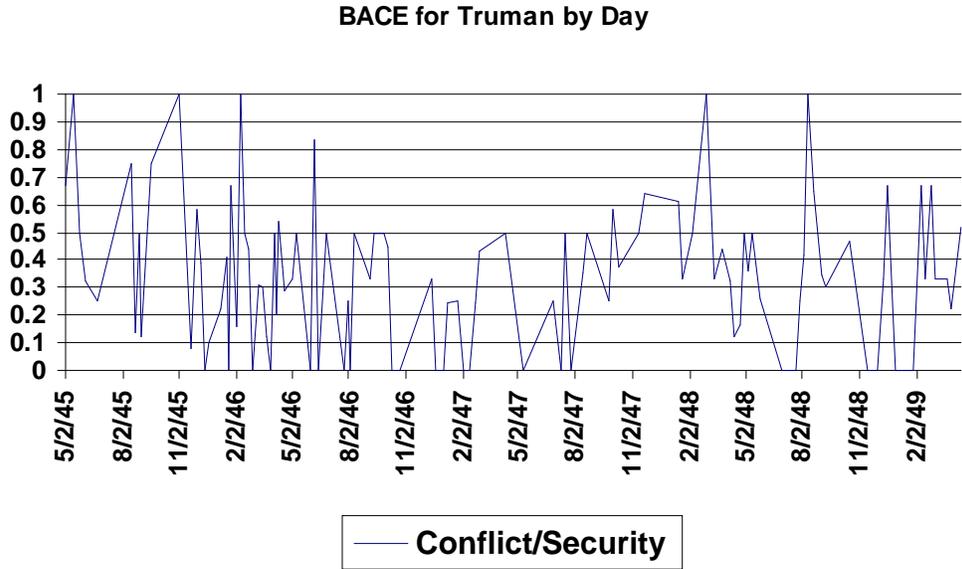


Figure 5.12



Analysis

Testing for trait stability across time at the day, quarter, and year units of aggregation has shown that relative consistency (as defined in Hypotheses 7 & 8) in trait scores for any leader does not exist. While we failed to find trait stability as defined by previous research (Hermann 1980), the combination topic-trait scores for each leader at the year level failed to show extreme variation from one period to the next. Moving to the quarter level of aggregation, extreme variation from one period to the next is only observed for some Presidents, and in many cases we observe what appears to be trends in scores with occasional spikes at certain points. Using day as the unit of aggregation we see tremendous variation from one day to the next for all Presidents on all trait-topic combinations.

These findings can be interpreted in a number of ways. Because we are dealing with a psychological construct that has no tangible substance that can be extracted and measured using physical instruments, there is no way to determine whether these score variations reflect changes in personality itself or if they measure something other than personality. Using the definition of personality taken in this study¹³, we would expect scores to remain at least relatively stable from one period to the next. Given that position, we find conflicting results depending on the time unit used for aggregation. Comparing Presidents from one year to the next, we find partial confirmation of stability as observed in the tests of Hypotheses 7 & 8. Moving to the quarter and then to the day units of aggregation, we find little to no stability (respectively).

This leaves us with at least three possibilities, first, that Leadership Trait Analysis is a worthless tool which simply produces random output that is not unique to any individual leaders. Second, the Leadership Trait Analysis technique is a complicated method for counting word and

¹³ “Personality refers to a construct that is introduced to account for the regularities in an individual’s behavior as he responds to diverse stimuli”. (Greenstein, 1969)

phrase preferences of individuals with no connection to politically relevant leadership characteristics. Finally, we are left with the possibility that, while not measuring personality per se, the technique does measure politically relevant characteristics of a leader's psychological makeup.

The complete answer to this question can only be found in future research, however, some preliminary points may be addressed using the data at hand. The first possibility, that the technique simply produces random scores that are completely independent of the leader can be discounted on two counts. First, the scores do not appear to be random spikes across time from point to point for all leaders – therefore the fact that some leaders show some stability for some periods at higher levels of aggregation, and the fact that we see patterns of moves from high to low on the same trait for different topics, leads us to believe that this is not the case. Second, the ANOVA tests performed in Chapters 3 and 4 which tested for trait score difference by President showed (with the exception of in-group bias) that the scores produced in this analysis differ significantly depending upon President. With this we can confidently say that the scores produced by the Leadership Trait Analysis technique are not simply random numbers without connection to the word and phrase usage of individual leaders.

While these results show that it is likely that the scores produced by the technique are connected to the leaders whose speech they are derived from, we are left without a method to determine whether they are connected to politically relevant components of a leader's psychological makeup. In order to assess whether or not the scores produced by the system are truly related to a leader's political behavior and not simply reflective of their linguistic patterns, we must correlate scores with observed behavior. In fact, the lack of temporal stability in trait scores provides us with ample opportunity to assess this situation. We may find that changes in

trait scores precede behavioral events or they may follow significant events. If they precede the events then the method would be useful in predicting future behavior. If not, then it may only be useful for ad hoc analyses of leader behavior and may be more reflective of the rhetoric used by leaders or reflect a change in the leader's approach to the world based on the events that occur in their environment – like Carter's change in worldview after the Soviet invasion of Afghanistan in 1979.

Conclusion

Temporal stability as previously conceived does not exist. Even when all data points from an entire year are aggregated into a single unit we fail to see rank classification stability in nearly two-thirds of the trait-topic combinations for the four Presidents. Looking at the Presidents in relation to one another, in no cases do all four Presidents remain in the same category across all four periods. We see even less consistency as we move to aggregation by quarter and finally we see no stability when aggregating by day. These tests, however, are strict tests that make it highly likely that stability would not be falsely identified. Both at the year and quarter aggregation units we do see some stability and the evidence of patterns across time. These findings, along with statistically significant difference in trait scores for the majority of traits by leader lead us to believe, at the very least, that the scores produced by Leadership Trait Analysis are not simply random numbers derived from the words spoken by these leaders. The implications for future research using the Leadership Trait Analysis technique will be discussed in greater detail in the conclusion to this study.

Chapter 6

Sample Size Requirements for Reliable Trait Score Creation

As discussed in Chapter 2, the lack of trait stability over topic and across time suggests that it is inappropriate to specify a sample size for the approximations of a leader's score on any leadership trait. Because the mean trait score for a leader moves across a significant range of values across time, using a sample of the total available documents for a leader will preclude the researcher from identifying these shifts. Examining a few hypothetical scenarios will illustrate this point. Suppose we have two leaders (Leader A and Leader B) each with 1000 spontaneous documents (split evenly across four years for 250 documents per year) available for coding. After coding all documents for each leader and aggregating at the year level, we find that Leader A's conceptual complexity scores remain quite stable with a score of .49 for years 1 and 3, and a score of .51 for years 2 and 4. With this outcome, averaging the data across the four years and obtaining a score of .50 for complexity would likely reflect the leader's central tendency on that trait. This presents us with a clear number by which we can compare the score generated by our random sample to assess its accuracy in approximating the true value obtained from coding all available documents. Leader B, on the other hand, shows significant variation across time on conceptual complexity in a highly patterned way. Aggregating at the year level, we find that at year 1 the conceptual complexity score is .2, at year 2 it moves to .4, at year 3 it moves to .6 and at year 4 it peaks at .9. If a 100 document sample was randomly selected to produce trait scores, what number would we use to assess the accuracy of the sample in its attempt at approximation of the entire population of documents? It would be possible to take the mean trait score across the entire set of documents, but doing so ignores significant shifts in trait scores across time and is likely not reflective of the leaders central tendency for that trait.

In order to ensure that this holds true with the actual data gathered in this analysis, this chapter will provide a brief demonstration of the consequences of producing samples using the 50 response size suggested by Hermann before the advent of automated coding (Hermann 1983, 1999). Ten randomly generated samples of 50 responses each will be drawn from each President. This process will be performed three times for each President, once for each response topic (conflict, economic, and political). Segmenting the data by topic allows us to assess the effect instability over time plays in affecting sample accuracy by controlling for variation in trait scores across topic. Because of this control for response topic there is an increase in the likelihood that sample accuracy will be obtained.

Table 6.1 presents the mean trait scores for ten samples of President Bush’s public (ranked by value, Table 6.2 displays the results by sample) utterances where the topic dealt with conflict/security issues. Each sample contains 50 responses randomly selected using a computerized random number generator so each of Bush’s 1,789 responses on conflict/security issues had an equal chance of being included in each sample.

Table 6.1 Ordered mean scores for President Bush’s traits using 50 response samples from Conflict responses

	BACE	CC	DIS	IGB	PWR	SC	TASK
Low	0.27	0.51	0.06	0.04	0.19	0.49	0.51
	0.31	0.57	0.08	0.04	0.20	0.51	0.53
	0.32	0.57	0.08	0.05	0.20	0.51	0.54
	0.32	0.59	0.09	0.06	0.21	0.52	0.54
	0.33	0.60	0.09	0.07	0.22	0.53	0.54
	0.33	0.61	0.12	0.08	0.22	0.56	0.55
	0.33	0.62	0.14	0.08	0.25	0.57	0.57
	0.34	0.62	0.15	0.08	0.29	0.58	0.58
High	0.36	0.66	0.19	0.11	0.29	0.59	0.59
	0.38	0.68	0.20	0.13	0.30	0.59	0.66
Difference from high to low	0.11	0.17	0.14	0.09	0.11	0.1	0.15
Difference / Standard Deviation	2.20	3.40	3.50	4.50	3.67	1.11	2.50

The data in Table 6.1 show that using 50 response samples produce considerable variation in trait scores. The smallest absolute range between high and low sample scores is .09 for in-group bias. Looking at the range divided by the standard deviation of the population mean (used for categorizing leaders relative to one another, as was done in previous chapters) we see that the highest and lowest in-group bias scores differ by 4.5 standard deviation units. The smallest variation in terms of standard deviations from the population mean is self confidence which varies by slightly more than one standard deviation. Figure 6.2, using the population mean and standard deviations, displays the rank category value for the trait score derived by each sample.

Table 6.2 Change in classifications for 50 response samples drawn from all Bush responses where topic is conflict

Bush (Conflict)	BACE (Mod.)	CC (Mod.)	DIS (Mod.)	IGB (Mod.)	PWR (Mod.)	SC (High)	TASK (Low)	Incorrect Categorization*
Sample 1	Mod.	Mod.	Mod.	Mod.	Low	Low	Mod.	3
Sample 2	Mod.	Mod.	High	Low	Mod.	Low	Low	3
Sample 3	Mod.	Mod.	Mod.	Low	High	High	Low	2
Sample 4	Mod.	Mod.	Mod.	High	Low	High	Low	2
Sample 5	Mod.	High	Mod.	Mod.	Mod.	High	Low	1
Sample 6	Mod.	Mod.	Mod.	Mod.	Low	High	Low	1
Sample 7	Mod.	Mod.	Mod.	High	High	Low	Low	3
Sample 8	Mod.	High	Mod.	Low	Mod.	Low	Low	3
Sample 9	Low	Mod.	Low	Low	Mod.	High	Low	3
Sample 10	Mod.	High	High	Mod.	High	Low	Low	4
Incorrect Categorization*	1	3	3	6	6	5	1	25

* when compared to the value obtained by coding all Bush conflict responses (listed in parentheses at each trait's column head)

The results clearly show that, depending on the sample, Bush looks like a different person. In four of the traits (conceptual complexity, distrust of others, in-group bias and need for power) the scores obtained from various samples span all three classifications – low, moderate and high. While belief in one's ability to control events mean scores across the ten samples predominantly fall into the moderate category, they span almost the entire range of that category – from .29 to .39. Task orientation mean scores predominantly fall into the low category, ranging from .51 to

.59 across nine samples. The sample scores for self confidence are evenly split with half falling in the moderate category and the other half falling into the high category. Looking at table 6.2 we see that across the ten samples, Bush is ranked differently 25 times for an average of 2.5 traits

One caution must be noted before accepting these results. The 50 response minimum suggested by Hermann was based on responses consisting of 100 words or more. The decision was made in this analysis to use all codeable material, regardless of word count. In order to assess what effect word count has on sample reliability, an additional set of 10 random samples were produced for each leader selecting only from responses consisting of 100 words or more. Tables 6.3 and 6.4 show the results for this second set of random samples.

Table 6.3 Mean scores for President Bush’s traits using 50 response samples (of at least 100 words)

	BACE	CC	DIS	IGB	PWR	SC	TASK
Low	0.27	0.61	0.06	0.03	0.18	0.54	0.54
	0.28	0.61	0.07	0.04	0.20	0.54	0.54
	0.30	0.62	0.09	0.05	0.21	0.54	0.56
	0.31	0.62	0.09	0.05	0.22	0.54	0.56
	0.31	0.64	0.09	0.06	0.22	0.55	0.59
	0.31	0.65	0.10	0.06	0.23	0.57	0.59
	0.32	0.66	0.13	0.07	0.23	0.57	0.60
	0.34	0.66	0.15	0.09	0.24	0.59	0.61
High	0.36	0.66	0.15	0.09	0.26	0.61	0.62
	0.38	0.67	0.16	0.09	0.27	0.61	0.64
Difference from high to low	0.11	0.06	0.10	0.06	0.09	0.07	0.10
Difference / Standard Deviation	2.20	1.20	2.50	3.00	3.00	0.78	1.67

Table 6.4 Change in classifications for 50 response samples drawn from all Bush responses where topic is conflict and word count is greater than 99

Bush (Conflict 99+ words)	BACE (Mod.)	CC (Mod.)	DIS (Mod.)	IGB (Mod.)	PWR (Mod.)	SC (High)	TASK (Low)	Incorrect Categorization*
Sample 1	Mod.	Low	Mod.	Mod.	Mod.	High	Low	1
Sample 2	Mod.	Low	Mod.	Mod.	Mod.	Mod.	Mod.	3
Sample 3	Mod.	Mod.	Mod.	Low	Mod.	Mod.	Mod.	3
Sample 4	Mod.	Mod.	High	Low	High	Mod.	Mod.	5
Sample 5	Mod.	Low	Mod.	Mod.	Mod.	High	Low	1
Sample 6	Low	Mod.	Mod.	Low	Low	High	Low	3
Sample 7	Mod.	Mod.	Mod.	Low	Mod.	High	Low	1
Sample 8	Mod.	Mod.	Low	Mod.	Low	Mod.	Low	3
Sample 9	Low	Low	Low	Low	Mod.	High	Low	4
Sample 10	Mod.	Mod.	Mod.	Low	High	High	Mod.	3
Incorrect Categorization*	2	4	3	6	4	4	4	27

* when compared to the value obtained by coding all Bush conflict responses (listed in parentheses at each trait's column head)

The data in table 6.4 show that there is a reduction in the range of scores on six of the seven traits when only using responses of 100 words or more, while the range across samples for belief in one's ability to control events remains the same. Looking at Table 6.4, however, shows that limiting the sample to responses consisting of 100 words or more actually leads to a slight increase in the number of incorrect category placements.

This simple exercise clearly reinforces the need to collect as much data as is available for any individual we are interested in studying. The instability of trait scores across time combined with small sample sizes leads to both contradictory and unreliable trait scores from one sample to the next. Depending on which of the 10 samples were used to assess Bush's leadership profile, he would appear to be quite a different person. The same test was performed for the other Presidents included in this study, with similar results in all cases with the exception of Truman for responses where the word count was greater than 99. This test could not be conducted for Truman's responses over 100 words in length because there were less than 50 responses for each

trait once the data was split by topic. Because Truman had so few spontaneous responses over 100 words in length, breaking our responses into smaller subsets based on response topic leaves us with three sets of less than 50 responses. Because we could not meet the 50 count threshold, this test was not conducted for Truman.

Chapter 7

Implications for the Future of Leadership Trait Analysis

The findings in the previous chapters of this study both support and undermine some of the key assumptions made by researchers employing the Leadership Trait Analysis technique to better understand the foreign policy process. By revisiting some critical questions asked of the approach in the past using a significantly more rigorous application of the technique over an incredible amount of data, this study sets the groundwork for future studies using the Leadership Trait Analysis method. Regardless of whether each specific finding support or undermines our previous understanding of the Leadership Trait Analysis technique, taken together they provide a better understanding of what needs to be done to further refine the approach and more effectively test its validity for explaining behavior in the international system. This concluding chapter first summarizes the findings in each of the preceding chapters and concludes with the implications of this study for future research employing the Leadership Trait Analysis technique.

Summary of Findings

Hypothesis 1: Trait scores derived from prepared utterances will vary significantly from trait scores derived from spontaneous utterances.

The data clearly confirm Hypothesis 1 with each President in this study exhibiting statistically significant variation by type of material used to produce trait scores for at least five of seven leadership traits. Both the self-confidence and task orientations produce statistically significant differences for all four Presidents while the remaining five traits produce statistically significant differences for three of the four leaders. Truman and Clinton had the least statistically significant trait totals at five traits each; Bush exhibited statistically significant variation on six traits while

Johnson showed statistically significant differences on all seven traits. In no cases did two leaders have the same trait appear statistically insignificant. In addition to examining whether statistically significant differences exist for the four Presidents, each of the President-trait combinations was also examined to see if spontaneity resulted in substantively different trait classifications. Using the rank classifications obtained through comparison to a larger sample of world leaders, eighteen of the twenty two traits which exhibited statistically significant variation also result in a shift of rank classifications depending on the type of material used to derive trait scores.

Hypothesis 2: Differences in trait scores derived from prepared versus spontaneous utterances are random.

An examination of the direction of change between trait scores derived from prepared versus spontaneous material fails to confirm Hypothesis 2. With the exception of distrust of others, within each trait all leaders who exhibited statistically significant variation between spontaneous and prepared material showed the same direction of change between the two types of material. Belief in one's own ability to control events, in-group bias, and need for power produce *lower* scores on spontaneous material than for prepared material. Conceptual complexity, self confidence, and task orientation produce *higher* scores on spontaneous material than for prepared material. While the magnitude of these differences was not consistent either across leader or across trait, a distinctive pattern was identified between the two types of material for six of the seven traits.

Hypothesis 3: President is a significant source of variation in trait scores across all data coded in this study.

Performing ANOVA on all spontaneous responses for the four leaders included in this study, we find that all but one trait (in-group bias) shows statistically significant variation by President. This finding generates added support for claims that the Leadership Trait Analysis technique is capable of identifying distinct scores for different leaders and is thus useful in comparing the different behaviors each leader initiates while in office.

Hypothesis 4: Mean trait scores for each President will exhibit statistically significant variation across response topic.

While response topic was not a significant source of variation for all traits for all leaders, 68% of the traits did show significant variation. Of the traits that did not have significant variation, there does not appear to be any discernible pattern as to why each trait for each leader was or was not significant – with the possible exception of in-group bias that exhibited insignificant variation for three of the four Presidents. Among Presidents, we see that Bush and Clinton both show variation in six of the seven traits, with in-group bias being insignificant for Bush and belief in one's ability to control events for Clinton. Of Johnson's seven personality traits only three traits showed significant variation based on response topic with belief in one's ability to control events, in-group bias, need for power, and self-confidence all failing to show statistically significant variation by response topic. Truman showed slightly more variation with four of the seven traits being significant (conceptual complexity, distrust of others, and in-group bias failed to show statistically significant variation based on response topic). Taken as a whole, this leaves only one trait where three Presidents failed to show significant variation (in-group bias), and one trait where two Presidents failed to show significant variation (belief in one's ability to control events). The remaining traits each had one President who failed to show significant variation by response topic.

Hypothesis 5: Trait scores which exhibit statistically significant variation based on response topic will also exhibit substantively significant variation.

Of the nineteen traits reporting statistically significant variation for the four Presidents, eleven also showed substantive variation between the three topics of interest deemed relevant to the study of foreign policy decision-making (conflict/security, economics, and politics). Closer examination of the substantive variation observed shows some pattern within traits, however it is questionable as to whether a pattern exists within President. For conceptual complexity (which showed statistically significant variation for three Presidents), all of the Presidents showed substantive variation, with economic responses scoring higher than at least one other category and never less than another category. Task orientation (which showed statistically significant variation for all four Presidents) showed substantive variation for only one President. Belief in one's ability to control events (where only two Presidents showed significant statistical variation) failed to show substantive variation in both cases. Distrust of others, need for power, and self-confidence each had three Presidents who showed statistical variation and in each case two Presidents showed substantive variation as well.

Within President the only pattern that is observed is that economic responses tend to be classified in a higher category than the political or conflict/security responses. Both conflict/security and political responses' relative comparison to the other two categories appear to fluctuate randomly. While we see substantive variation for nine of the twenty-eight leader-trait combinations measured in this study, we are left with little ability to predict when that variation will occur for any leader and for any specific trait. These results are at the very least a partial confirmation of Hypothesis 5, and the remainder of this study's analysis is conducted separately for each response topic.

Hypothesis 6a: President is a significant source of variation in trait scores across all conflict/security responses coded in this study.

Hypothesis 6b: President is a significant source of variation in trait scores across all economic responses coded in this study.

Hypothesis 6c: President is a significant source of variation in trait scores across all political responses coded in this study.

Performing ANOVA on all responses within each topic we find that the majority of traits show significant variation by leader for the three response topics assessed in this study. The only trait that did not show statistically significant variation by President for all three topics was in-group bias, which adds further support to the argument that in-group bias is not measuring anything distinct across individuals. Further research, however, must be conducted to assess whether in-group bias' failure to show statistically significant differences in this analysis is a product of a) the small sample of leaders included in this study or b) neglecting to perform in-group assignments for the in-group candidates identified by Profiler Plus discussed earlier.

For conflict and political responses, all of the remaining traits show statistically significant variation by President. For economic responses we fail to see statistically significant difference on two additional traits. In addition to in-group bias, need for power and self-confidence fail to show significant variation ($p < .05$). This leaves us with seventeen of twenty-one trait-topic combinations showing statistically significant variation by President.

Hypothesis 7: Mean trait scores for each time period will be classified in the same category ranking.

Hypothesis 8: A President's rank for any given trait (compared to the other three Presidents in the study) will remain constant across the entire four years.

Testing for trait stability across time at the day, quarter, and year units of aggregation has shown that relative consistency in trait scores for any leader does not exist. While we failed to find trait stability as defined by previous research (Hermann 1980), the combination topic-trait scores for each leader at the year level failed to show extreme variation from one period to the next. Moving to the quarter level of aggregation, extreme variation from one period to the next is only observed for some Presidents, and in many cases we observe what appears to be trends in scores with occasional spikes at certain points. Using day as the unit of aggregation we see tremendous variation from one day to the next for all Presidents on all trait-topic combinations.

Hypothesis 9: Trait scores based on 50 responses constitute a reliable sample of material to construct a personality profile.

The findings in Chapters 4 and 5 regarding instability in trait scores across topic and time led to the expectation that Hypothesis 9 would fail to be confirmed. The instability of trait scores across time combined with small sample sizes leads to both contradictory and unreliable trait scores from one sample to the next. Depending on which of the 10 samples were used to assess Bush's leadership profile, he would appear to be quite a different person. The same test was performed for the other Presidents included in this study, with similar results in all cases. Because previous research was based on the 100 word count threshold for response selection, two sets of samples were generated with the second set limited to responses of 100 words or more. While the range of scores in these samples decreased, the trait classification volatility still remained.

Implications for future research

The dramatic reduction in costs associated with coding leaders afforded by automated content analysis systems such as Profiler Plus will undoubtedly revolutionize the study of foreign policy decision-making. It is no longer inconceivable to conduct studies involving hundreds of leaders to test the role individuals play in the formulation of their nation's foreign policy. While confirming some of the initial assumptions used by the Leadership Trait Analysis approach, this study has also found reason to be cautious before conducting such ambitious studies. Four major implications can be identified from the data produced in this analysis. The type of material used for trait assessment; the contextual nature of the scores produced by the coding system; the need to link scores with behavior; and a reconsideration of the way trait scores are calculated, especially when examining trait scores across time.

What material should be used to assess an individual's leadership traits?

The existence of both statistical and substantive differences between prepared and spontaneous statements further supports the assumption that trait scores derived from prepared material is influenced by factors such as speechwriters and attempts at conveying a specific message to an audience (Hermann, etc....). This finding contradicts Brian Dille's study of the prepared and spontaneous utterances of George Bush, Sr. and Ronald Reagan which failed to find significant differences between the conceptual complexity scores for the two types of material. Because Dille's study also incorporated an automated process for coding similar to that used in this analysis, the contradictory findings are not likely a function of coding reliability. The most likely cause for this discrepancy is the amount of material used to produce conceptual complexity scores in Dille's study. His study was limited to 15 speeches and 15 press

conferences for each leader, while this study incorporated over 3000 speeches and 1000 press conferences across the four Presidents.

Depending on the question being asked by the researcher, it is advisable to limit the type of documents used in the development of trait scores. For approaches like Leadership Trait Analysis, which are interested in assessing the personal characteristics of a specific individual, the findings show that use of prepared materials is likely to lead to biased results. These approaches are advised to make use of spontaneous utterances such as those found in press conferences and interviews given by the leader in question. Other approaches (such as Operational Code analysis and Image Theory), however, are interested in the policy of an entire administration or government body and thus would benefit from the inclusion of all available material that can be attributed to members of those groups. This can include speeches as well as policy statements that are not necessarily spoken by a person in a public setting, but provide insight into the policy stance taken by such groups.

While these findings lend substantial support to Hermann's position regarding the influence of speechwriters and rhetoric affects on prepared material, the question still remains as to whether the scores of spontaneous material are reflective of the "true" personality of the leader being examined. In order to more fully explore this issue, additional research is suggested to further delineate the effect of spontaneity on trait scores. This study followed the approach of previous research (Dille 2000; Dille and Young 2000) by dividing documents into two categories: speeches and interviews/press conferences. Having prepared all of the documents for each President found in the Public Papers of the Presidents series, a number of additional subcategories can be identified that are being subsumed within the prepared category. It can be argued that speeches made at campaign rallies, which contain a focused and fairly consistent

message across multiple documents and time periods may differ significantly from speeches made from the Oval Office during a national crisis. Campaign speeches, addresses to professional organizations, state of the union addresses, all serve different purposes and it may be that some of those types of documents are more reflective of the leader's personality. If this is found to be true, limiting studies to spontaneous material may unnecessarily restrict data available for trait score production.

An additional category can also be identified which has yet to be coded in the existing research: private material. The 25-year limit on classified documents has passed for many of the modern U.S. Presidents and thus affords the opportunity to assess spoken words that were delivered with an expectation of privacy. Such material is less susceptible to claims of rhetoric and instrumentality in speech. The possibility does exist that the leader may be attempting to win over competing factions within the private group, and thus still employing instrumental speech, given enough material over different settings and topics this could be assessed to test its validity. Additional research testing whether a significant difference exists between scores produced from private material and public spontaneous statements can be used to further assess the likelihood that spontaneous material is more reflective of the leader's "true" personality. If a significant difference can not be found, it would lend credibility to the claim that spontaneous material is sufficient for assessment of a leader's personal characteristics and that instrumental speech is not a significant concern. If, on the other hand, a significant difference is found between the two types of material, it could call into question the likelihood that spontaneous material is a true reflection of a leader.

Contextual Nature of the Expressed Personality Measured by Leadership Trait Analysis

Significant debate has occurred over the past decades regarding the nature of what Leadership Trait Analysis purports to be measuring. In the early stages of the research program, traits were conceived to be static representations of the core attributes of an individual's personality (Hermann 1980 and earlier...). Critics such as Rasler, Thompson, and Chester (1980) argued that the results produced by the system were not stable across time and called into question the usefulness of the technique. With the evolution of the research program, the traits are now understood to be reflective of a "person-in-context", measuring the interplay between the individual and the context in which that individual is placed. The more recent studies employing the Leadership Trait Analysis technique acknowledge that leaders do not exist in a vacuum and that certain characteristics (both internal and external to the leader) will affect the "expressed" personality as measured by the coding system (Hermann 1999, and others).

As with the other questions examined in this study, previous research was hampered by the inability to code large volumes of data and thus more reliably assess the stability of trait scores across time. The findings of this study confirm that the "person-in-context" approach more accurately reflects the scores produced by the coding system. The first contextual variable addressed in this analysis, response topic, was found to be a valid source of differentiation between trait scores for individual leaders. The findings show that a leader's trait scores show a measurable difference depending on the topic addressed in their response to the questions asked of them in interviews and press conferences.

Determining that differences in trait scores based on response topic exist has significant implications for studies that use a single trait score (calculated by the average of all responses coded) to determine the leader's trait description. This suggests that a leader's expressed

personality will differ depending on the topic of the questions they are asked during press conferences. Calculating a single value to assess the leader's central tendency for any given trait will therefore be influenced by the distribution of topics within the material used to generate the trait scores. A leader who is asked primarily conflict/security questions during their press conferences will exhibit different overall trait scores than if that same person had been asked primarily economic questions. The magnitude and direction of such a change, however, is not predictable based on the findings of Chapter 4 and thus can not be used to adjust a leader's overall score based on the distribution of responses of varying topics. Depending on the specific research question asked by a researcher using the Leadership Trait Analysis technique, this finding leads to different implications.

Researchers interested in using a leader's central tendency on a trait as a predictor of behavior must take into account the makeup of the data used to produce trait scores. Studies such as Preston's research on the role a leader's conceptual complexity and need for power has on the structure of their advisory system will need to do one of two things. If the assumption is made that a leader's central tendency on a trait is reflective of the average of various relevant topics, then scores must be calculated giving equal weight to the various topics included within the material used to generate trait scores.

The other option is to assess whether the behavior in question falls discretely within one of the response topics measured during the coding process. If a researcher is interested in how a leader will behave in economic negotiations with a trading partner, the average value of trait scores derived solely from economic topics will likely be a more appropriate measure of how the leader will behave than the average of their mean scores produced through analysis of responses on political and conflict/security topics. Research on advisory groups, on the other hand, will

most likely need to rely on an average of all topics because the advisory group is responsible for a variety of policy issues (including economic and defense policies) and may reflect the leader's expressed personality on more than just political issues.

While some research questions will best be answered using a leader's central tendency for each trait, other studies can make use of the shifts in scores throughout different time periods, the other contextual factor used to assess stability in trait scores. While the early research on Leadership Trait Analysis expected to find stability in trait scores across all contexts and over time, as noted previously the research program has evolved over the past decades to account for the effect of context on a leader's trait scores measured by the coding system. The findings in Chapter 5 confirm the validity of this perspective, based on the score fluctuations found across all of the time aggregations (year, quarter, and day) used in this analysis.

While a visual inspection of the fluctuation of trait scores over time show that trait scores show varying degrees of stability (or instability) over time, this study did not attempt to determine whether these changes were a product of time in office itself or if other contextual factors are being masked within it. It can be hypothesized that time in office does explain variation in trait scores, the most obvious explanation being that leader's learn more about their new role over time and change accordingly. Although additional research focused explicitly on this issue is required, given the data produced in this analysis it is unlikely that the changes across time are a function of learning effects. A review of the figures representing trait-topic combinations for the four leaders over time (see Appendix B for a complete collection of all figures for trait-topic combinations aggregated by year and quarter) shows mixed support for learning across an administration. The majority of the trait-topic combinations for each President show both positive and negative adjustments across time as opposed to continual positive or

continual negative shifts from one period to the next. Figure 7.1 provides an example of such shifts, displaying conceptual complexity scores for response topics for each President. With the exception of Clinton's score remaining constant for the first two years in office, each data point for each President moves in the opposite direction as the previous data point's movement. This does, however, define learning as a linear progression in a certain direction over time. If learning is defined as a general trend across the entire time period, then Bush's scores over the four years displayed in Figure 7.1 could be seen as following a general trend of reduced conceptual complexity over time.

Figure 7.1

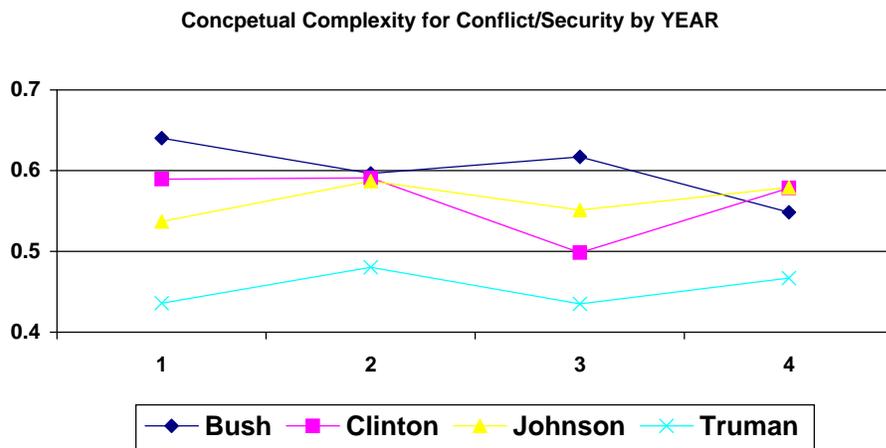


Figure 7.2, displays need for power scores based on conflict responses, and provides further support for such a hypothesis, with both Presidents Bush and Clinton showing an initial decrease in scores followed by an increase in the last quarter. President Johnson shows a steady rise in scores from year 1 to year 3 with just a slight drop in year 4.

Figure 7.2

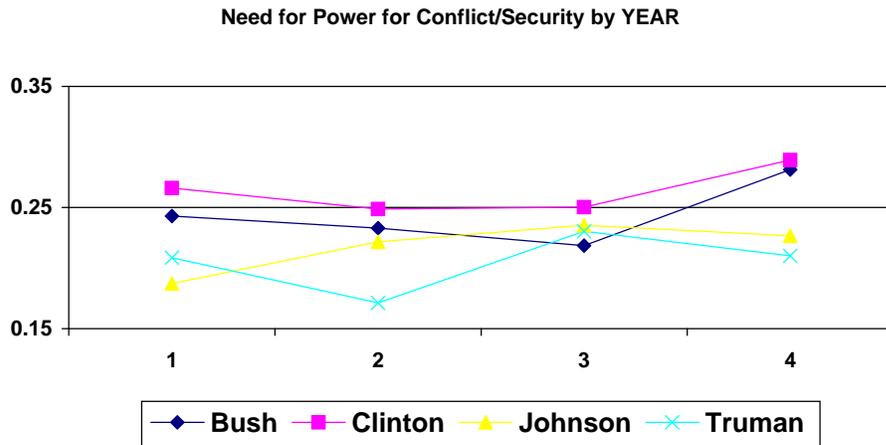
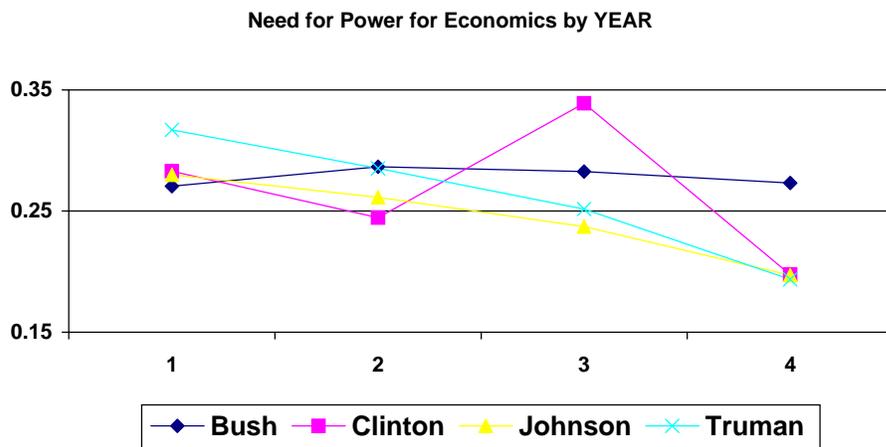


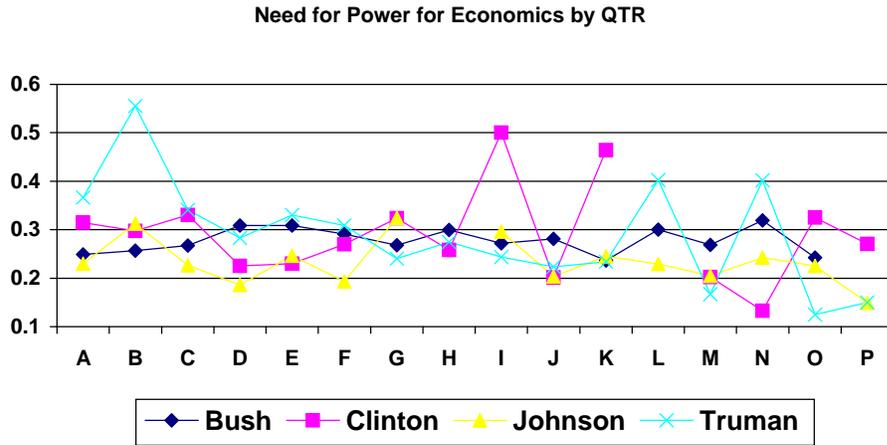
Figure 7.3, displays need for power for economic responses, provides what can be construed as clear support for a learning effect for both Johnson and Truman. Clinton, on the other hand scores alternate between increases and decreases while Bush remains relatively stable.

Figure 7.3



While year to year trends are useful for broad generalizations, they provide very few data points to assess shifts across time. Looking at data aggregated at the quarter level we have more opportunities to assess whether a learning effect is occurring. Figure 7.4 displays the same trait-topic combination as Figure 7.3 but uses the quarter as the unit of aggregation.

Figure 7.4



As was argued in the chapter on trait stability over time, aggregating at smaller units of time results in less stability from period to period (as would be expected). At this time aggregation we can still see a general decrease in scores for Truman across the entire four years, however Johnson's score increases in five of the fifteen periods following the start date and decreases in nine periods, with no valid data coded for economic responses in one period (H). Looking through Appendix B, it becomes evident that the data aggregated at the quarter level provides little support for a continuous learning trend across the majority of time periods.

In addition to learning, other time in office factors may support time as a contextual variable of interest. Regularly occurring events such as elections may account for changes over time. Determining what effect regularized events and learning play in trait score variation over time is important in order to determine what other contextual factors influence trait scores. If such events combined with learning fail to account for the majority of fluctuation across time, other factors will need to be identified and tested to better understand the nature of such change. The outbreak of war (or an impending conflict), an economic recession, legislative defeat on a

bill the President actively supported, or a shift in the control of the Congress – each of these events which may or may not occur during a President’s administration is a potential source of trait score variation from one time period to the next. In addition to these discrete events, prolonged influences such as growing criticism of the President over time may lead to patterned shifts in trait scores (unfortunately for most Presidents, this is a relatively common experience – Truman with the Korean war, Johnson with the Vietnam war, Nixon with Watergate, and Clinton with the Lewinsky scandal to name just a few).

In addition to events that occur during a President’s administration, other contextual factors exist which may explain trait score fluctuations. While the audience to whom the President is speaking remains relatively consistent across the spontaneous material examined in this analysis (the majority of the documents are press conferences), occasionally the President fields questions at speaking engagements for different segments of the constituency. It can be hypothesized that a leader’s conceptual complexity increases when speaking to audiences with more education. Task orientation may increase when speaking to government employees where the President is interested in increasing support for his administration’s policies. A variety of similar hypotheses could be developed and tested for a whole host of contextual factors.

Much of the previous research employing the Leadership Trait Analysis technique has focused on a leader’s sensitivity to context to predict whether a leader’s trait scores are more susceptible to change with differing contexts (Hermann 1977, 1983, 1984, 1987, 1997, 1999; Preston 1996, 2001). Hermann argues that determining a leader’s self-other orientation provides insight into whether the leader is sensitive to their context. This orientation is calculated by comparing the leader’s conceptual complexity and self confidence scores to one another. If a leader’s conceptual complexity score is higher than the self confidence score the leader is,

(among other things) “... sensitive to situational cues and act based on what they sense is acceptable under current conditions” (Hermann 1999, p. 18). Leaders whose conceptual complexity score is lower than their self confidence score will be less sensitive to context, being “fairly unresponsive or insensitive to cues from the environment. Instead they reinterpret the environment to fit their view of the world” (p. 18). Accordingly, we would expect leaders whose conceptual complexity score is higher than their self confidence score to show less stability in trait scores over time while those whose self confidence scores are higher than their conceptual complexity scores will likely exhibit increased stability across time. If the scores are relatively equal, Hermann suggests examining the leader’s relative score rankings through comparison with other world leaders. Table 7.1 reports conceptual complexity and self-confidence scores (and category rankings) derived from spontaneous material for each President.

Table 7.1 Conceptual Complexity and Self-Confidence scores

President	Conceptual Complexity Score	Self Confidence Score	Conceptual Complexity Ranking	Self Confidence Ranking
Bush	.63	.55	High	High
Clinton	.59	.58	Moderate	High
Johnson	.58	.56	Moderate	High
Truman	.49	.53	Low	High

Presidents Bush and Truman show clear differences in their scores for the two traits while Clinton and Johnson’s scores are relatively equal. Bush’s conceptual complexity score is higher than his self confidence score, which would suggest he is highly sensitive to his context. Truman, showing the reverse comparison, would appear to be less sensitive to his context. However, looking at each leader’s trait scores over time, we see the reverse effect. While both leader’s show instability, Truman shows substantially more spikes in trait scores from one period to the

next than does Bush. Assessing Clinton and Johnson are difficult because their conceptual complexity ranks moderate when compared to other world leaders. The manual for Leadership Trait Analysis provides predictions regarding stability only if conceptual complexity and self confidence both rank high or both rank low, but does not provide predictions when the scores are either moderate or if one score is high and the other low. This preliminary assessment of the sensitivity to context approach suggests that more research be conducted to assess whether the outcome is a product of using only four leaders for assessment or if relative data amounts influence the findings. While the total number of coded responses for Truman far exceed that used in previous studies, Bush's total coded words and responses far outweigh that for Truman. A more thorough analysis of additional leaders is necessary before making claims regarding the validity of the sensitivity to context technique employed in the Leadership Trait Analysis research program.

Because the findings from this study rarely display linear progression across all time periods, it is likely that multiple contextual factors are influencing changes in trait scores across time. In order to determine the relative importance of such factors, additional research should be conducted which attempts to isolate similar contextual settings among different leaders and assess the trait scores for each leader during that context. Observed variance in trait scores across different leaders when contextual factors are held constant would lend additional support to the findings in this study that show trait score variation by leader. Given additional resources, looking at all post-war U.S. Presidents over multiple terms in office may uncover additional data regarding election cycle patterns. Such studies will undoubtedly add to the understanding of individual differences and allow researchers to better assess the implications of a substantial increase or decrease in trait scores from one period to the next. Without identifying whether

additional contextual factors exist that significantly affect the expressed personality of leaders, what appears to be a significant shift in a leadership trait may simply be a reflection of the environment in which the leader is found. The vast amounts of material available for U.S. Presidents combined with their relatively similar contextual experiences (such as those identified above) provides researchers interested in these questions the opportunity to better answer such questions.

Expanding the Linkage Between Trait Scores and Behavior

While previous research has identified general correlations between trait scores and certain behaviors, the inability to code significant amounts of data for a large population of leaders has prevented the establishment of direct linkages between trait scores and specific behaviors in the foreign policy realm. Does a .3 score in distrust of others equate with genocidal behavior? Does a 25% increase in task orientation over a 3 month period correspond with mobilization for war? Now that the capacity to code data for every world leader for whom we can collect documents is possible, we are able to create complex models which allow us to take into account leader personality and their environmental constraints in order to explain past events and predict future ones.

Given the nature of the findings in previous chapters, this is an essential task. To date, our analysis of trait scores have been limited to comparisons to the population's mean score. Without a clearer understanding of the relationship between trait scores and specific behaviors, however, this provides an unsatisfactory method for distinguishing between different leaders. By the nature of the calculation, the leaders who are placed in the high and low categories make up 28% of the population of leaders in the entire comparison sample. This leaves 62% of the population in the Moderate category. This poses problems for use as a measure of trait stability

both across topic and across time. When examining whether trait scores changed in a substantive way in Chapters 3-5, trait score classifications were used to assess whether a President shifted between classifications across document type, topic, or time. Why this is problematic becomes evident with a hypothetical scenario. Assume that there are two leaders (A and B) and through measurement of trait scores across time we find that Leader A's scores for conceptual complexity move between .55 and .58 across the entire period of analysis while Leader B's scores move between .6 and .8. Depending on the placement of the threshold for determining what score constitutes a "High" ranking, Leader A may appear to be much less stable than Leader B. If the threshold is .56, Leader A will shift between moderate or high depending on the period in question. Leader B, on the other hand, would appear completely stable by being ranked high throughout the entire period of analysis.

Future research examining this issue can focus on a few things. First, can behaviors be identified that correspond to specific trait scores by leaders? The lack of stability in trait scores over time will actually increase our ability to determine whether certain scores are correlated with specific behaviors by providing us with multiple data points for each leader. We may find that dropping below a certain score threshold for a trait leads to the decision to engage in a pre-emptive strike against an adversary. If such connections can be identified, the ability to predict specific events would be possible. In addition to absolute scores, relative score shifts may also be connected to behavior. Previous research (Suedfeld, Tetlock, and Ramirez, 1977) has identified that downward shifts in conceptual complexity occur in public rhetoric of leaders as far as six months in advance of the outbreak of war. Previous studies which relied on central tendency scores may find that deviations from the norm (such as the use of alternative advisory systems for the resolution of specific situations, may correlate with certain shifts in scores across time).

Another possibility may be that certain behaviors group around certain trait score intervals. Rather than being a linear trend, we may find that specific behaviors are compartmentalized within discrete ranges in trait scores, leading to a stepping effect.

The benefits of a closer examination of the potential linkages between trait scores and actual behavior are significant and will most likely be some of the most interesting work derived from this research program. To engage in such research, however, it will be critical to first identify the additional variables (if any exist) which influence leader trait scores. This study has already shown that material type, response topic, and time have significant effect on trait scores. Without determining the extent to which such situationally dependent variables influence scores, the connection between such scores and behaviors may be spurious.

Sampling and Measurement Issues

For researchers interested in using central tendency trait scores for leaders, the findings of this study make clear, with no reservations, that trait scores derived from samples consisting of fifty spontaneous responses are not sufficiently reliable or accurate when compared to the scores derived from all publicly available data. This result is not surprising given the variation in trait scores observed by both topic and across time. That initial research employing the Leadership Trait Analysis technique found this number to be sufficient is most likely the product of the limitations brought on by hand coding of documents which made systematic tests of randomly selected samples an extremely costly endeavor. With the ability to code all available documents afforded by automated content analysis, issues related to sample size need not be a concern. While there are additional costs associated with collection and preparation of all available documents, these are by far inconsequential compared to the added reliability and validity in trait scores produced when coding the entire set of materials available for a leader. Assuming that the

inconsistent magnitude and direction changes in scores across topic and time found in this study are representative of the population of world leaders as a whole, these differences (and their magnitudes) will only be identified through the coding of the entire set of documents available for a leader. Exceptions to this may include studies that seek to examine behavior in a discrete time period (the Cuban Missile Crisis for example) and trait score data before and after the event are not a part of the analysis. One note of caution against employing such an approach is that as of yet the research has not identified whether absolute trait scores or changes in scores over time are better predictors of a leader's actions. If all data for the leader's entire term in office (and preferably data prior to entering office as well) were available it would likely place the trait scores of the time period of interest in the larger context of the leader as a whole.

During the examination of the sampling data it became apparent that using small response samples not only affects attempts at approximating the central tendency of a trait score over an entire leader's term in office, but also assessments of trait stability over time. When moving to the aggregation of trait scores by day as opposed to quarter, the number of extreme scores (zeroes and ones) for each trait increase significantly. This is, in part, a function of splitting the data by response topic. During some press conferences (especially brief press availabilities made when foreign dignitaries visit Washington) there may be few to no questions related to economic or conflict/security concerns. Table 7.2 provides details regarding the frequency of responses by day for George Bush, Sr.'s conflict data. Once the data is split by response topic, nearly 20% of the days in Bush's trait over time analysis are based on a single response for conflict issues, nearly 30% for economic issues, and slightly more than 10% for political issues.

Table 7.2 Response Frequency by day for George Bush, Sr.

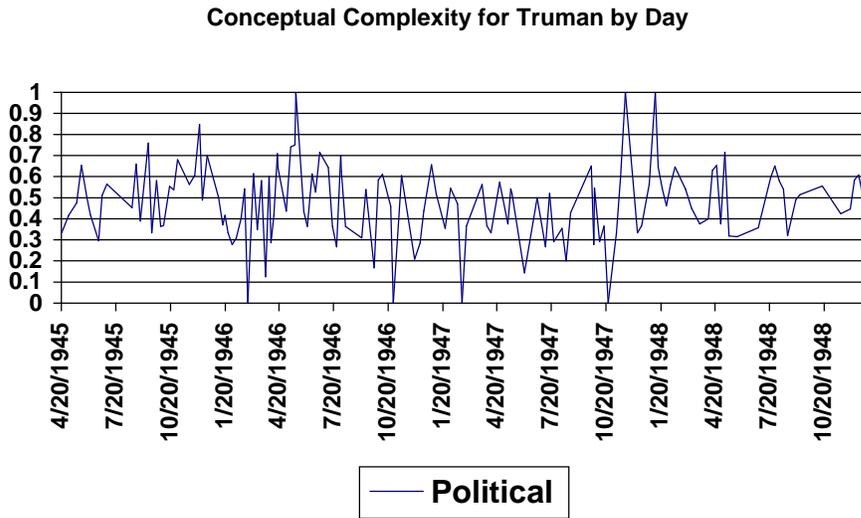
# of Responses	# of days for Conflict responses (% of total: 294)	# of days for Econ. responses (% of total: 226)	# of days for Pol. responses (% of total: 362)	# of days for All responses (% of total: 407)
1	56 (19%)	63 (28%)	42 (12%)	14 (3%)
< 5	136 (46%)	148 (66%)	150 (41%)	71 (18%)
< 10	224 (76%)	197 (87%)	227 (63%)	165 (41%)
< 25	281 (96%)	225 (99%)	334 (92%)	310 (76%)
< 50	293 (99%)	226 (100%)	359 (99%)	394 (97%)

These results led to a further examination of the data and the way scores are measured. As discussed in Chapter 1, the coding scheme developed by Hermann instructs the coder to look for certain words and phrases that indicate the absence or presence of the seven leadership traits. If the word or phrase identified reflects the presence of the trait, a score of 1 is given to that text. If the word or phrase identified reflects the absence of the trait, a score of 0 is given to that text. All scores within a response are then averaged to produce a score between 1 and 0 for the entire response. All responses are then aggregated to produce the overall trait score for the leader. Calculating scores in this method assumes that responses provide equal opportunities for coding decisions, which proves to be inaccurate. Coding decisions where a 1 or 0 is applied to the text within responses do not occur uniformly throughout text. Giving each response equal weight in the calculation of scores fails to take into account two things. First, it either fails to recognize or considers unimportant the issue of word count across various responses. With the utilization of automated content analysis, the 100 word count limit was removed so responses can be as little as 1 word (answering “yes” or “no”) or as many as the leader chooses to speak in response to a question (Bush’s longest single response consists of nearly 1300 words). It should be noted, however, that in research preceding automated analysis, the 100 word count limit coupled with user selection of responses to code most likely dramatically reduced the range in words across

the responses used for trait score production. The second, and more important, item that this method of calculating trait scores fails to consider is that coding opportunities are not a direct function of word count. Depending on the trait, a response consisting of 100 words may have as little as no coding decisions (no words or phrases indicating either the presence or absence of a trait) to as many as 100 (while theoretically possible – the leader could repeat the word “always” 100 times and the conceptual complexity measure would code 100 zero scores – the number would undoubtedly be lower). Because leaders do not always speak using words that the coding system recognizes as containing the presence or absence of a trait, “coding opportunities” (defined as a score of 1 or 0 for any word or phrase in the material being coded) differ across responses.

This method of calculation further places into question data aggregated at the day level of analysis. Days where the response count for a specific topic is as low as one may also contain only one coding decision (the minimum necessary for a score to be present – responses with no coding opportunities are considered missing data and are not included in score calculation). Because of the binary nature of the coding system, fewer coding opportunities in a response present a greater the likelihood that an extreme score (a 1 or 0) will be obtained. On days where the trait score is a product of a single response with only one coding opportunity, the probability that a 1 or 0 is obtained is 100%. A preliminary test of the affect coding opportunity has on trait scores was performed for Truman’s conceptual complexity where the response topic was politics. Figure 7.6 shows Truman’s conceptual complexity scores across time aggregated at the day for all data included in this study.

Figure 7.6 Conceptual Complexity Scores for Truman Aggregated by Day



As expected, days exist where the trait scores reach either 1 or 0, although the total number of days with such scores is only seven (of 148 days where a conceptual complexity score was calculated) or 5% of the time. This may not seem very significant until the data in Figures 7.7 and 7.8 are examined. Figure 7.7 excludes days where there were less than ten coding opportunities (where the mathematical probability of a 1 or 0 score is 0.2%) while Figure 7.8 excludes those days where there were less than twenty-five coding opportunities (where the mathematical probability of a 1 or 0 score is .000006%). Note that the data is not calculated differently, it simply excludes dates with less coding opportunities than the established threshold.

Figure 7.7 Conceptual Complexity Scores for Truman, Aggregated by Day (excluding days with less than 10 coding opportunities)

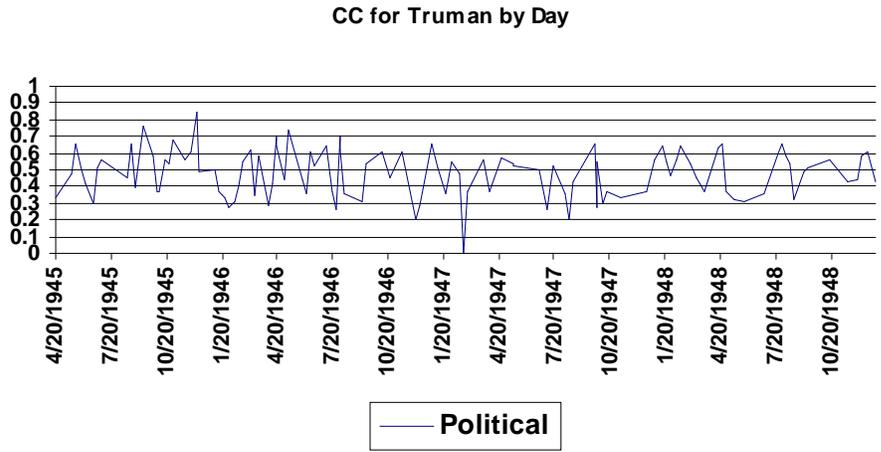


Figure 7.8 Conceptual Complexity Scores for Truman, Aggregated by Day (excluding days with less than 25 total opportunities)

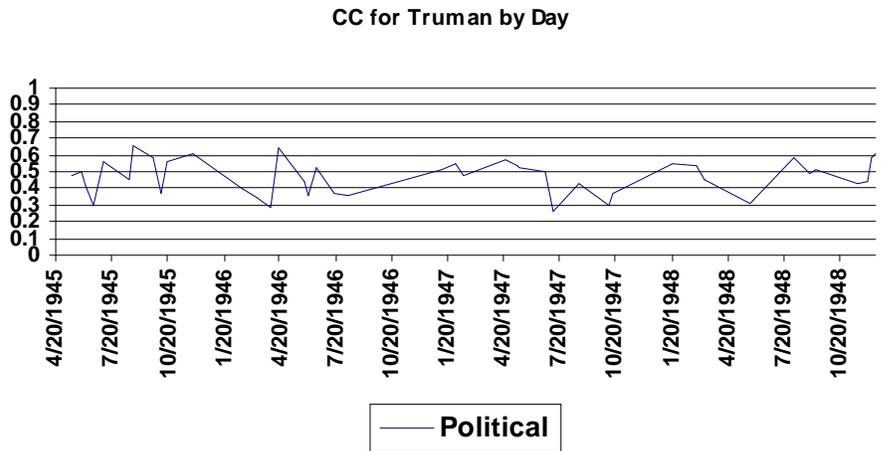
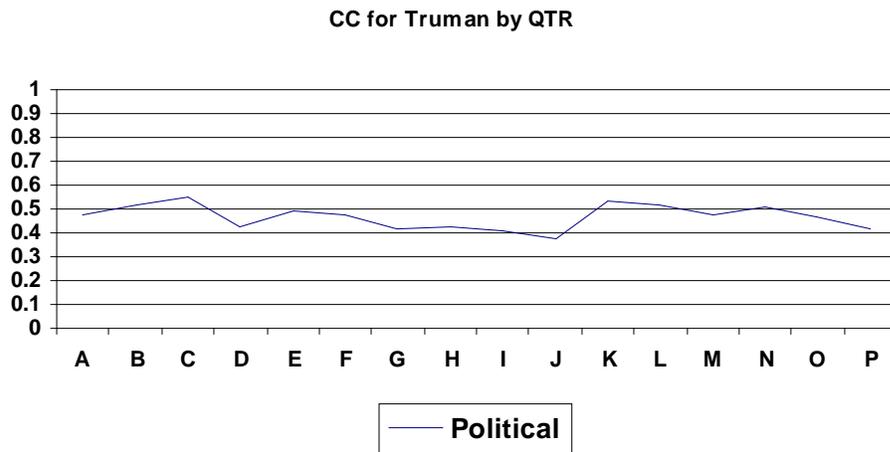


Figure 7.9 Conceptual Complexity Scores for Truman Aggregated by Quarter



Looking at Figure 7.7 (days with at least 10 coding opportunities), it becomes immediately noticeable that the number of days with the most extreme score of 1 or 0 drops to a single day. In addition, using the 10 coding opportunity threshold decreases the observed range of scores by .15 (to .85). However, the most noticeable difference appears in Figure 7.8 (days with at least twenty-five coding opportunities). Not only does the remaining 0 value disappear, but the observed range of scores drops to .39, significantly less than the full range of values observed when all data is used. Especially interesting is the comparison between the 25 coding opportunity threshold data and that of Figure 7.4, Truman's conceptual complexity data for political topics aggregated at the quarter level. While the data does not track identically with one another, given that Figure 7.3 is still being aggregated at the day level shows remarkable stability and comparison to the data aggregated at three month intervals. This clearly shows that minimal coding opportunities are a significant source of the extreme variation observed at the day level of aggregation.

It is important to note that using the thresholds in this manner does result in data being excluded from the analysis. Employing the 10 coding opportunity threshold results in the exclusion of 25% (36 of 144) days where data on conceptual complexity is available for

Truman's political responses. Employing the 25 coding opportunity threshold results in 71% of the days where data on conceptual complexity for Truman's political responses being excluded.

While it can be shown that the number of coding opportunities significantly influences the amount of variation in trait scores across time, there is no clear answer regarding what qualifies as an appropriate threshold for the minimum number of coding opportunities necessary for confidence in trait scores produced by the coding system. Whatever choice researchers make will have implications for both the reliability and the utility of studies employing such thresholds. If the threshold is set at a relatively high level, take the 1200-1500 respondents used in polling the U.S. electorate as a comparison, the confidence that the score reflects the leader's "true" personality will increase. At the same time, however, in order to get that many coding opportunities (segmented across response topic because it was found to be a significant source of variation) the unit of aggregation may move from a single response to six month's worth of responses, resulting in a dramatic decrease in the ability to measure sensitivity in scores during a crisis or other events which do not occur over months or years.

An examination of the average coding opportunities per 5,000 words for the population of leaders provided by Social Science Automation shows that a significant amount of data is necessary to reach that threshold. Table 7.2 reports the average coding opportunities for 5,000 words and the high and low ranges. For some leaders gathering that many English language spontaneous uttered words is impossible. Not only would it eliminate many individuals who scholars are interested in studying (Saddam Hussein, Kim Jong-Il, Hu Jintao, Osama Bin Laden to name just a few) it would also severely restrict the ability to assess trait score differences across time. The average number of words collected for the world leaders who make up the comparison base is slightly more than 50,000 words. Using the average number of coding

opportunities per 5000 words, we would be left with only one or two periods for analysis, depending on the trait. This would severely limit our ability to test for contextual factors as well as use the technique to predict behavior in crises or other situations where decisions are made at a single day or a single week. Also, if the data is spread across a decade, we may find that the period of time required to collect 1500 coding opportunities varies across the leader's total available data. This would add additional difficulties to the assessment of multiple leaders for comparison purposes.

Table 7.2 Coding Opportunities for World Leaders in Comparison Population

Leadership Trait	Average Coding Opportunities per 5000 words	Lowest Number of Coding Opportunity per 5000 words	Highest Number of Coding Opportunity per 5000 words	Range
BACE	132	86	184	98
CC	297	259	339	80
DIS	141	107	197	90
IGB	113	74	156	82
PWR	125	82	175	93
SC	75	23	147	124
TASK	202	168	246	78

On the other hand, if a much smaller threshold is used, such as the 10 or 25 coding opportunity used to restrict data in Figures 7.2 and 7.3, many of the issues listed above are no longer a concern. Whether time periods with less than the threshold are excluded or a time period is selected which guarantees the threshold is achieved, all world leaders for whom we have spontaneous material would provide sufficient material to test the various hypotheses regarding trait stability recommended in this study. The downside to such an approach is that while we would be able to perform such tests, questions regarding the validity of scores at this limited threshold would remain. If coding opportunities are parallel to respondents to a poll in terms of

measuring a leader's "true" personality, than we should be highly skeptical that such material accurately reflects the person-in-context we wish to use for behavioral explanations.

One way to determine what the appropriate threshold should be is to collect data for a variety of leaders over a variety of periods during their tenure in office where trait scores can be linked to the leader's behavior (which assumes that the behavioral linkage sought above does actually exist). Producing scores for a variety of situations where coding opportunities vary, we can then match the predictions to observed behavior and see at what point a satisfactory level of predictive ability is achieved. Until that is accomplished, it is imperative that researchers employing the Leadership Trait Analysis technique refrain from aggregating at the response level and, at the very least, take into consideration the number of coding opportunities present when examining trait stability over time.

Conclusion

The ability to code more data than could ever have been imagined just ten years ago affords political psychologists an awesome opportunity to better understand the leaders who shape the foreign policy of countries around the world. This study has started that process through the examination of four US Presidents and asked critical questions of one technique used to study decision-making. While providing more complete answers to some fundamental questions that have faced the literature over the past decades, it has also expanded the questions that need to be asked by future researchers employing both the Leadership Trait Analysis technique and other remote assessment methodologies. While the desire to rush forward and code all available data for the latest leader of interest may be tempting, caution should be used before relying on assumptions that have yet to be validated through studies of larger samples of data. This study has identified a number of areas of future research which will allow us to further

refine the techniques used to study world leaders. The potential to dramatically improve our understanding of the role leaders play in international affairs is here, we need only take advantage of labor saving techniques such as automated content analysis to produce rigorous studies which take into account the role played by a leader's environmental context. This study has taken a first step towards performing such analyses by examining some of the fundamental assumptions of the Leadership Trait Analysis technique.

I believe the next step in this research agenda is to expand this study to the remaining modern presidents. While coding all verbal utterances of four U.S. president's first term in office exceeds by far anything that has been done in the past, we are still dealing with a sample of only four leaders. Extending the sample to the eleven presidents since Franklin Delano Roosevelt will provide more confidence in assessing whether or not the results of this study are peculiar to the four presidents chosen. In addition to increasing the number of presidents, coding all documents for multiple terms in office for those presidents who served multiple terms will offer additional insight into the trait score variations observed over time. While time and labor restraints in data collection and markup prevented their use in this analysis, I believe it is important to expand future research to include all verbal statements made by this group of presidents for their entire time in office.

In addition to expanding the scope of the study to include additional presidents, the other pressing need for the Leadership Trait Analysis research program (as well as other decision making approaches) is to more solidly connect such assessments to actual leader behavior. While this study was able to perform statistical calculations on numbers produced by a content analysis system, the real value of this technique is to explain and predict the behavior of world leaders. While tests of statistical significance may show that a difference exists between different types of

material or across different response topics, we are left without a satisfactory measure for substantive significance. Until the connection to behavior is made, we are left with the analysis of numbers which may or may not correlate with the phenomena of interest to scholars in this field: the foreign policy behavior or world leaders. This study has set the stage for future research by determining that trait scores do not, as was previously thought, remain stable across time. It also found that leader's scores for the leadership traits measured by the technique vary depending upon the topic the leader is talking about. While time itself may be a variable that explains trait score variation, it most likely is a surrogate for other events or characteristics present in a leader's environment. Identifying and controlling for such factors is a necessary first step in making this connection to behavior.

While the road to a more sophisticated use of decision making techniques is likely to be long, the advent of automated content analysis schemes has made the road accessible. Without such technology the field would be forever plagued with questions of inter-coder reliability and small sample sizes. Now that those obstacles have either been eliminated or severely reduced, the foreign policy decision making field will allow us to better understand the role specific individuals play in the events which shape the lives of every person on this planet.

Bibliography

- Alison, Graham and Philip Zelikow. *Essence of Decision: Explaining the Cuban Missile Crisis*. Second edition. New York: Longman, 1999.
- Allport, Gordon W. *Pattern and Growth in Personality*. New York: Holt, Rinehart and Winston, 1937.
- Allport, Gordon W. *The Nature of Personality: Selected Papers*. Cambridge: Addison-Wesley, 1950.
- Axelrod, Robert. *Structure of Decision*. Princeton: Princeton University Press, 1976.
- Baldwin, David A. "Neoliberalism, Neorealism, and World Politics". In David A. Baldwin, ed., *Neorealism and Neoliberalism: The Contemporary Debate*. New York: Columbia University Press, 1993.
- Barber, James D. *The Presidential Character: Predicting Performance in the White House*. Third edition, Englewood Cliffs, New Jersey: Prentice-Hall, 1985.
- Buzan, Barry, Charles Jones and Richard Little. *The Logic of Anarchy: Neorealism to Structural Realism*. New York: Columbia University Press, 1993.
- Cottam, Martha L. *Foreign Policy Decision Making: The Influence of Cognition*. Boulder: Westview Press, 1986.
- Cottam, Martha L. "The Carter Administration's Policy Toward Nicaragua: Images, Goals, and Tactics." *Political Science Quarterly* 107 (1992a): 123-46.
- Cottam, Martha L. "Contending Dramas in American Foreign Policy." In Martha Cottam and C. Shi, eds., *Contending Dramas: A Cognitive Approach to International Organization*. New York: Praeger, 1992b.
- Cottam, Martha L. *Images and Intervention: U.S. Policies in Latin America*. Pittsburgh: University of Pittsburgh Press, 1994.
- Cottam, Richard. *Foreign Policy Motivation*. Pittsburgh: University of Pittsburgh Press, 1977.
- Craik, Kenneth H., Robert Hogan, and Raymond N. Wolfe.
- Dille, Brian. "The Prepared and Spontaneous Remarks of Presidents Reagan and Bush: A Validity Comparison for At-a-Distance Measurements." *Political Psychology* 21 (2000): 573-85.
- Dille, Brian and Michael Young. "The Conceptual Complexity of Presidents Carter and Clinton: An Automated Content Analysis of Temporal Stability and Source Bias." *Political Psychology* 21 (2000): 587-96.

- Dougherty, James E. and Robert L. Pfaltzgraff, Jr. *Contending Theories of International Relations: A Comprehensive Survey*. Third edition. New York: Harper and Row, 1990.
- Ewen, Robert. *An Introduction to Theories of Personality*. Third edition, Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.
- Forgus, Ronald and Bernard Shulman. *Personality: A Cognitive View*. Englewood Cliffs, New Jersey: Prentice Hall, 1979.
- George, Alexander L. and Juliette L. George. *Woodrow Wilson and Colonel House: A Personality Study*. New York: Dover 1964.
- George, Alexander L. "The 'Operational Code': A Neglected Approach to the Study of Political Leaders and Decision Making." *International Studies Quarterly* 13, (1969): 190-222.
- George, Alexander L. "The Causal Nexus Between Cognitive Beliefs and Decision-Making Behavior: The 'Operational Code' Belief System." In Lawrence S. Falkowski, ed., *Psychological Models in International Politics*. Boulder: Westview Press, 1979.
- George, Alexander L. *Bridging the Gap: Theory and Practice in Foreign Policy*. Washington D.C.: United States Institute of Peace Press, 1993.
- George, Alexander L. and Juliette L. George. *Presidential Personality and Performance*. Boulder: Westview Press, 1998.
- Greenstein, Fred I. *Personality and Politics: Problems of Evidence, Inference, and Conceptualization*. Chicago: Markham, 1969.
- Greenstein, Fred I. "Eisenhower's Leadership Style". In Gunther Bischoff and Stephen Ambrose, eds., *Eisenhower*. Baton Rouge: Louisiana State University Press, 1995.
- Hermann, Margaret G. "Some Personal Characteristics Related to Foreign Aid Voting of Congressmen." In Margaret G. Hermann, ed., *A Psychological Examination of Political Leaders*. New York: Free Press, 1977.
- Hermann, Margaret G. "Explaining Foreign Policy Behavior Using Personal Characteristics of Political Leaders." *International Studies Quarterly* 24 (1980a): 7-46.
- Hermann, Margaret G. "Comments on Foreign Policy Makers' Personality Attributes and Interviews: A Note on Reliability Procedures." *International Studies Quarterly* 24 (1980b): 67-73.
- Hermann, Margaret G. *Handbook for Assessing Personal Characteristics and Foreign Policy Orientations of Political Leaders*. Columbus: Mershon Center Occasional Papers, 1983a.

- Hermann, Margaret G. "Assessing Personality-at-a-Distance: A Profile of Ronald Reagan." *Mershon Center Quarterly Report*, vol. 7., Columbus: 1983b.
- Hermann, Margaret G. "Personality and Foreign Policy Decision Making: A Study of 53 Heads of Government." In D.A. Sylvan and S. Chan, eds., *Foreign Policy Decision-Making: Perceptions, Cognition, and Artificial Intelligence*, pp. 53-80. New York: Praeger Press, 1984.
- Hermann, Margaret G. "Ingredients of Leadership." In M.G. Hermann, ed., *Political Psychology: Contemporary Problems and Issues*, pp. 167-92. San Francisco: Jossey-Bass, 1986.
- Hermann, Margaret G. "Leaders' Foreign Policy Orientations and the Quality of Foreign Policy Decisions." In S. Walker, ed., *Role Theory and Foreign Policy Analysis*, pp. 161-98. Durham: Duke University Press, 1987a.
- Hermann, Margaret G. "Assessing the Foreign Policy Role Orientations of Sub-Saharan African Leaders." In S. Walker, ed., *Role Theory and Foreign Policy Analysis*, pp. 161-98. Durham: Duke University Press, 1987b.
- Hermann, Margaret G. "Defining the Bush Presidential Style." *Mershon Center Memo*. Columbus: Ohio State University Press, 1989.
- Hermann, Margaret G. and Thomas Preston. "Presidents, Advisers, and Foreign Policy: The Effect of Leadership Style on Executive Arrangements." *Political Psychology*, 15. no. 1 (1994a): 75-96.
- Hermann, Margaret G. and Thomas Preston. "Presidents and Their Advisers: Leadership Style, Advisory Systems, and Foreign Policymaking." In E. Wittkopf, ed., *The Domestic Sources of American Foreign Policy: Insights and Evidence*, pp. 340-56.
- Hermann, Margaret G. "Advice and Advisers in the Clinton Presidency: The Impact of Leadership Style." In S. Renshon, ed., *The Clinton Presidency: Campaigning, Governing, and the Psychology of Leadership*, pp. 149-64. Boulder: Westview Press, 1995.
- Hermann, Margaret G. and Thomas Preston. "Presidents, Leadership Style, and the Advisory Process. In E.R. Wittkopf and J. M. McCormick, eds., *The Domestic Sources of American Foreign Policy: Insights and Evidence*, pp. 351-68. Lanham Maryland: Rowman and Littlefield, 1999.
- Hermann, Margaret G., Thomas Preston, Baghat Korany, and Timothy M. Shaw. "Who Leads Matters: The Effects of Powerful Individuals." In Joe D. Hagan and Margaret G. Hermann, eds., *Leaders, Groups, and Coalitions: Understanding the People and Processes in Foreign Policymaking*, pp. 83-132. Malden, Massachusetts: Blackwell Publishing, 2001.

- Hollis, Martin and Steve Smith. *Explaining and Understanding International Relations*. Oxford: Clarendon Press, 1991.
- Holsti, Ole. "Cognitive Dynamics and Images of the Enemy: Dulles and Russia." In David Finlay, Ole Holsti, and Richard Fagen, eds., *Enemies in Politics*. Chicago: Rand McNally, 1967.
- Holsti, Ole. "The 'Operational Code' as an Approach to the Study of Political Leaders: John Foster Dulles' Philosophical and Instrumental Beliefs." *Canadian Journal of Political Science* 3 (1970): 123-57.
- Hudson, Valerie M. "Culture and Foreign Policy: Developing a Research Agenda." In Valerie M. Hudson, ed., *Culture and Foreign Policy*. Boulder: Lynne Reiner, 1997.
- Janis, Irving L. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Revised edition, Boston: Houghton Mifflin, 1982.
- Jervis, Robert. *The Logic of Images in International Relations*. Princeton: Princeton University Press, 1970.
- Jervis, Robert. *Perception and Misperception in International Politics*. Princeton: Princeton University Press, 1976.
- Kaarbo, Juliet. "Prime Minister Leadership Styles in Foreign Policy Decision Making: A Framework for Research." *Political Psychology* 18 (1997): 533-81.
- Kaarbo, Juliet and Margaret G. Hermann. "Leadership Styles of Prime Ministers: How Individual Differences Affect the Foreign Policy Process." *Leadership Quarterly* 9, no. 3 (1998): 243-63.
- Keohane, Robert O. "Realism, Neorealism, and the Study of World Politics". Robert O. Keohane, ed., *Neorealism and its Critics*. New York: Columbia University Press, 1986.
- Kille, Kent. *Leadership and Influence in the United Nations: A Comparative Analysis of the Secretaries-General*. Doctoral Dissertation, Columbus: Ohio State University, 2000.
- Khong, Yuen Foong. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton: Princeton University Press, 1992.
- Leites, Nathan. *The Operational Code of the Politburo*. New York: McGraw-Hill, 1951.
- Leites, Nathan. *A Study of Bolshevism*. New York: Free Press, 1953.

- Marfleet, B. Gregory. "The Operational Code of John F. Kennedy During the Cuban Missile Crisis: A Comparison of Public and Private Rhetoric." *Political Psychology* 21 (2000): 545-58.
- Morgenthau, Hans J. *Politics Among Nations*. New York: Knopf, 1948.
- Miller, George. "The Magic Number Seven, Plus or Minus Two". *Psychological Review* 63 (1956): 81-97.
- Neustadt, Richard E. *Presidential Power and the Modern Presidents: The Politics of Leadership From Roosevelt to Reagan*. New York: The Free Press, 1990.
- Preston, Thomas. *The President and His Inner Circle*. New York: Columbia University Press, 2001.
- Rasler, Karen A., William R. Thompson, and Kathleen Chester. "Foreign Policy Makers, Personality Attributes, and Interviews: A Note on Reliability Problems." *International Studies Quarterly* 24 (1980): 47-66.
- Renshon, Stanley A. *Psychological Needs and Political Behavior: A Theory of Personality and Political Efficacy*. New York: The Free Press, 1974.
- Ripley, Brian. "Psychology, Foreign Policy, and International Relations Theory." *Political Psychology* 14 (1993): 533-81.
- Schafer, Mark. "Issues in Assessing Psychological Characteristics at a Distance: An Introduction to the Symposium". *Political Psychology* 21 (2000): 511-527.
- Schafer, Mark and Scott Crichlow. "Bill Clinton's Operational Code: Assessing Source Material Bias." *Political Psychology* 21 (2000): 559-571.
- Schafer, Mark and Michael D. Young. "Is There Method in our Madness? Ways of Assessing Cognition in International Relations." *Mershon International Studies Review* 42 (1998): 63-96.
- Shimko, Keith. *Images and Arms Control: Perceptions of the Soviet Union in the Reagan Administration*. Ann Arbor: University of Michigan Press, 1991.
- Skowronek, Stephen. "Presidential Leadership in Political Time." In Michael Nelson, ed., *The Presidency and the Political System*. Fourth edition. Washington D.C.: Congressional Quarterly Press, 1995.
- Snyder, Richard, H.W. Bruck and Burton Sapin. *Foreign Policy Decision-Making: An Approach to the Study of International Politics*. New York: Free Press, 1954.

- Suedfeld, Peter, Philip E. Tetlock, and C. Ramirez. "War, Peace, and Integrative Complexity: U.N. Speeches on the Middle East Problem." *Journal of Conflict Resolution* 21 (1977): 427-42.
- Taysi, Tanyel and Thomas Preston. "The Personality and Leadership Style of President Khatami: Implications for the Future of Iranian Political Reform." In Ofer Feldman and Linda O. Valenty (eds.), *Profiling Political Leaders: A Cross-Cultural Studies of Personality and Behavior*. Westport, Connecticut: Praeger, 2001
- Tetlock, Philip E. and Charles B. McGuire, Jr. "Cognitive Perspectives on Foreign Policy." In Samuel Long, ed., *Political Behavior Annual*, Boulder: Westview Press, 1985.
- Thucydides. *The History of the Peloponnesian War*. London: Oxford University Press, 1960.
- Vertzberger, Yaacov V.I. *The World in Their Minds: Information Processing, Cognition, and Perception in Foreign Policy Decisionmaking*. Stanford: Stanford University Press, 1990.
- Walker, Stephen G. "The Interplay Between Beliefs and Behavior: Henry Kissinger's Operational Code and the Vietnam war." *Journal of Conflict Resolution* 11 (1977): 129-68.
- Walker, Stephen G., and Timothy Murphy. "The Utility of the Operational Code in Political Forecasting." *Political Psychology* 3 (1981-1982): 24-60.
- Walker, Stephen G. "The Motivational Foundations of Political Belief Systems: A Re-Analysis of the Operational Code Construct." *International Studies Quarterly* 27 (1983): 179-201.
- Walker, Stephen G. "The Evolution of Operational Code Analysis". *Political Psychology* 11 (1990): 403-18.
- Walker, Stephen G., and Mark Schafer. "The Political Universe of Lyndon B. Johnson and His Advisors: Diagnostic and Strategic Propensities in their Operational Codes." *Political Psychology* 21 (2000): 529-543.
- Walker, Stephen G., Mark Schafer and Michael Young. "Operational Codes and Role Identities: Measuring and Modeling Jimmy Carter's Operational Code." *International Studies Quarterly* 42 (1998): 173-88.
- Waltz, Kenneth. *Man, the State, and War: A Theoretical Analysis*. New York: Columbia University Press, 1959.
- Waltz, Kenneth. *Theory of International Politics*. New York: McGraw Hill, 1979.
- Waltz, Kenneth. "Anarchic Orders and Balances of Power." In G. John Ikenberry, ed., *American Foreign Policy: Theoretical Essays*. Second edition. New York: Harper Collins, 1996.

Winter, David G., Margaret G. Hermann, Walter Weintraub, and Stephen G. Walker. "The Personalities of Bush and Gorbachev Measured at a Distance: Procedures, Portraits, and Policy." In G. John Ikenberry, ed., *American Foreign Policy: Theoretical Essays*. Second edition. New York: Harper Collins, 1996.

Young, Michael D. "Cognitive Mapping Meets Semantic Networks." *Journal of Conflict Resolution* 40 (1996): 395-414.

APPENDIX A

Data Tables for Trait Score Variation Across Time

Table A.1 Bush for Conflict by Year

Trait	Year	Response Count	Mean	Classification
BACE	1	297	0.3	Moderate
BACE	2	817	0.29	Low
BACE	3	518	0.3	Moderate
BACE	4	157	0.38	Moderate
CC	1	337	0.64	High
CC	2	872	0.6	Moderate
CC	3	536	0.62	Moderate
CC	4	177	0.55	Moderate
DIS	1	234	0.16	High
DIS	2	679	0.09	Moderate
DIS	3	460	0.13	Moderate
DIS	4	129	0.07	Low
IGB	1	212	0.03	Low
IGB	2	520	0.06	Moderate
IGB	3	405	0.06	Moderate
IGB	4	124	0.09	Moderate
PWR	1	296	0.24	Moderate
PWR	2	815	0.23	Moderate
PWR	3	518	0.22	Moderate
PWR	4	158	0.28	High
SC	1	291	0.54	Moderate
SC	2	800	0.58	High
SC	3	495	0.57	High
SC	4	146	0.52	Moderate
TASK	1	278	0.58	Low
TASK	2	756	0.58	Low
TASK	3	494	0.58	Low
TASK	4	163	0.59	Moderate

Table A.2 Bush for Econ by Year

Trait	Year	Response Count	Mean	Classification
BACE	1989	123	0.32	Moderate
BACE	1990	270	0.33	Moderate
BACE	1991	314	0.36	Moderate
BACE	1992	192	0.33	Moderate
CC	1989	126	0.66	High
CC	1990	270	0.67	High
CC	1991	325	0.66	High
CC	1992	205	0.64	Moderate

DIS	1989	102	0.11	Moderate
DIS	1990	224	0.1	Moderate
DIS	1991	278	0.1	Moderate
DIS	1992	178	0.08	Low
IGB	1989	90	0.11	High
IGB	1990	204	0.06	Moderate
IGB	1991	254	0.06	Moderate
IGB	1992	141	0.05	Low
PWR	1989	122	0.27	High
PWR	1990	270	0.29	High
PWR	1991	312	0.28	High
PWR	1992	189	0.27	High
SC	1989	120	0.58	High
SC	1990	257	0.56	High
SC	1991	303	0.61	High
SC	1992	188	0.6	High
TASK	1989	112	0.68	Moderate
TASK	1990	265	0.7	Moderate
TASK	1991	309	0.62	Moderate
TASK	1992	182	0.66	Moderate

Table A.3 Bush for Politics by Year

Trait	Year	Response Count	Mean	Classification
BACE	1989	1034	0.32	Moderate
BACE	1990	409	0.29	Low
BACE	1991	755	0.32	Moderate
BACE	1992	735	0.29	Moderate
CC	1989	1099	0.65	High
CC	1990	418	0.64	High
CC	1991	797	0.63	Moderate
CC	1992	744	0.64	Moderate
DIS	1989	851	0.11	Moderate
DIS	1990	340	0.07	Low
DIS	1991	676	0.1	Moderate
DIS	1992	651	0.11	Moderate
IGB	1989	661	0.07	Moderate
IGB	1990	277	0.07	Moderate
IGB	1991	512	0.05	Low
IGB	1992	534	0.08	Moderate
PWR	1989	1024	0.24	Moderate
PWR	1990	403	0.24	Moderate
PWR	1991	753	0.25	Moderate
PWR	1992	732	0.26	High
SC	1989	1014	0.55	Moderate
SC	1990	397	0.53	Moderate
SC	1991	733	0.54	Moderate
SC	1992	729	0.55	High

TASK	1989	1026	0.63	Moderate
TASK	1990	371	0.66	Moderate
TASK	1991	710	0.57	Low
TASK	1992	673	0.6	Moderate

Table A.4 Clinton for Conflict by Year

Trait	Year	Response Count	Mean	Classification
BACE	1993	549	0.35	Moderate
BACE	1994	805	0.36	Moderate
BACE	1995	75	0.3	Moderate
BACE	1996	194	0.39	High
CC	1993	558	0.59	Moderate
CC	1994	800	0.59	Moderate
CC	1995	70	0.5	Low
CC	1996	195	0.58	Moderate
DIS	1993	471	0.08	Moderate
DIS	1994	660	0.07	Low
DIS	1995	64	0.06	Low
DIS	1996	160	0.13	Moderate
IGB	1993	453	0.05	Low
IGB	1994	699	0.05	Low
IGB	1995	62	0.07	Moderate
IGB	1996	171	0.06	Low
PWR	1993	545	0.27	High
PWR	1994	800	0.25	Moderate
PWR	1995	75	0.25	Moderate
PWR	1996	193	0.29	High
SC	1993	513	0.62	High
SC	1994	731	0.64	High
SC	1995	72	0.61	High
SC	1996	175	0.6	High
TASK	1993	542	0.64	Moderate
TASK	1994	776	0.63	Moderate
TASK	1995	68	0.62	Moderate
TASK	1996	190	0.6	Moderate

Table A.5 Clinton for Economic by Year

Trait	Year	Response Count	Mean	Classification
BACE	1993	426	0.36	Moderate
BACE	1994	348	0.34	Moderate
BACE	1995	6	0.3	Moderate
BACE	1996	32	0.34	Moderate
CC	1993	433	0.66	High
CC	1994	363	0.65	High
CC	1995	6	0.47	Low
CC	1996	31	0.66	High

DIS	1993	388	0.1	Moderate
DIS	1994	310	0.08	Moderate
DIS	1995	5	0.08	Low
DIS	1996	26	0.02	Low
IGB	1993	373	0.06	Moderate
IGB	1994	328	0.04	Low
IGB	1995	6	0.03	Low
IGB	1996	30	0.13	High
PWR	1993	424	0.28	High
PWR	1994	348	0.24	Moderate
PWR	1995	6	0.34	High
PWR	1996	32	0.2	Low
SC	1993	393	0.62	High
SC	1994	326	0.57	High
SC	1995	6	0.8	High
SC	1996	28	0.71	High
TASK	1993	421	0.67	Moderate
TASK	1994	354	0.65	Moderate
TASK	1995	6	0.56	Low
TASK	1996	31	0.64	Moderate

Table A.6 Clinton for Political by Year

Trait	Year	Response Count	Mean	Classification
BACE	1993	356	0.37	Moderate
BACE	1994	542	0.37	Moderate
BACE	1995	77	0.34	Moderate
BACE	1996	160	0.34	Moderate
CC	1993	359	0.59	Moderate
CC	1994	549	0.56	Moderate
CC	1995	72	0.61	Moderate
CC	1996	167	0.54	Low
DIS	1993	323	0.1	Moderate
DIS	1994	469	0.1	Moderate
DIS	1995	60	0.1	Moderate
DIS	1996	140	0.09	Moderate
IGB	1993	250	0.08	Moderate
IGB	1994	432	0.06	Moderate
IGB	1995	57	0.03	Low
IGB	1996	115	0.05	Low
PWR	1993	351	0.28	High
PWR	1994	540	0.26	Moderate
PWR	1995	76	0.25	Moderate
PWR	1996	157	0.24	Moderate
SC	1993	348	0.56	High
SC	1994	527	0.56	High
SC	1995	77	0.55	Moderate
SC	1996	149	0.58	High

TASK	1993	333	0.62	Moderate
TASK	1994	529	0.62	Moderate
TASK	1995	67	0.61	Moderate
TASK	1996	154	0.62	Moderate

Table A.7 Johnson for Conflict by Year

Trait	Year	Response Count	Mean	Classification
BACE	1964	110	0.35	Moderate
BACE	1965	79	0.31	Moderate
BACE	1966	127	0.31	Moderate
BACE	1967	169	0.36	Moderate
CC	1964	110	0.54	Low
CC	1965	76	0.59	Moderate
CC	1966	125	0.55	Moderate
CC	1967	157	0.58	Moderate
DIS	1964	77	0.07	Low
DIS	1965	52	0.06	Low
DIS	1966	84	0.14	Moderate
DIS	1967	107	0.11	Moderate
IGB	1964	79	0.1	Moderate
IGB	1965	69	0.09	Moderate
IGB	1966	93	0.04	Low
IGB	1967	128	0.05	Low
PWR	1964	109	0.19	Low
PWR	1965	79	0.22	Moderate
PWR	1966	126	0.24	Moderate
PWR	1967	168	0.23	Moderate
SC	1964	100	0.57	High
SC	1965	71	0.63	High
SC	1966	118	0.6	High
SC	1967	155	0.55	Moderate
TASK	1964	106	0.7	Moderate
TASK	1965	75	0.62	Moderate
TASK	1966	113	0.63	Moderate
TASK	1967	138	0.63	Moderate

Table A.8 Johnson for Economic by Year

Trait	Year	Response Count	Mean	Classification
BACE	1964	90	0.37	Moderate
BACE	1965	29	0.42	High
BACE	1966	136	0.37	Moderate
BACE	1967	104	0.37	Moderate
CC	1964	87	0.63	Moderate
CC	1965	28	0.63	Moderate
CC	1966	131	0.64	High
CC	1967	104	0.68	High

DIS	1964	62	0.06	Low
DIS	1965	21	0.05	Low
DIS	1966	87	0.06	Low
DIS	1967	76	0.07	Low
IGB	1964	75	0.06	Moderate
IGB	1965	23	0.06	Moderate
IGB	1966	112	0.03	Low
IGB	1967	87	0.03	Low
PWR	1964	90	0.28	High
PWR	1965	29	0.26	High
PWR	1966	135	0.24	Moderate
PWR	1967	104	0.2	Low
SC	1964	72	0.54	Moderate
SC	1965	29	0.51	Moderate
SC	1966	115	0.56	High
SC	1967	88	0.62	High
TASK	1964	83	0.68	Moderate
TASK	1965	29	0.68	Moderate
TASK	1966	114	0.71	High
TASK	1967	94	0.71	Moderate

Table A.9 Johnson for Political by Year

Trait	Year	Response Count	Mean	Classification
BACE	1964	340	0.34	Moderate
BACE	1965	80	0.33	Moderate
BACE	1966	181	0.34	Moderate
BACE	1967	131	0.3	Moderate
CC	1964	303	0.53	Low
CC	1965	82	0.56	Moderate
CC	1966	172	0.58	Moderate
CC	1967	118	0.57	Moderate
DIS	1964	225	0.06	Low
DIS	1965	61	0.02	Low
DIS	1966	146	0.04	Low
DIS	1967	94	0.04	Low
IGB	1964	173	0.06	Moderate
IGB	1965	51	0.06	Moderate
IGB	1966	128	0.06	Low
IGB	1967	72	0.06	Low
PWR	1964	337	0.21	Moderate
PWR	1965	78	0.25	Moderate
PWR	1966	181	0.23	Moderate
PWR	1967	130	0.21	Moderate
SC	1964	319	0.56	High
SC	1965	76	0.53	Moderate
SC	1966	169	0.56	High
SC	1967	128	0.54	Moderate

TASK	1964	273	0.62	Moderate
TASK	1965	75	0.55	Low
TASK	1966	157	0.68	Moderate
TASK	1967	102	0.65	Moderate

Table A.10 Truman for Conflict by Year

Trait	Year	Response Count	Mean	Classification
BACE	1945	175	0.32	Moderate
BACE	1946	70	0.27	Low
BACE	1947	68	0.38	Moderate
BACE	1948	109	0.33	Moderate
CC	1945	150	0.44	Low
CC	1946	48	0.48	Low
CC	1947	63	0.44	Low
CC	1948	98	0.47	Low
DIS	1945	118	0.01	Low
DIS	1946	40	0.03	Low
DIS	1947	44	0.01	Low
DIS	1948	66	0.04	Low
IGB	1945	55	0.03	Low
IGB	1946	14	0.02	Low
IGB	1947	27	0.1	High
IGB	1948	26	0.1	Moderate
PWR	1945	166	0.21	Moderate
PWR	1946	63	0.17	Low
PWR	1947	60	0.23	Moderate
PWR	1948	100	0.21	Moderate
SC	1945	161	0.57	High
SC	1946	74	0.58	High
SC	1947	61	0.53	Moderate
SC	1948	106	0.59	High
TASK	1945	121	0.74	High
TASK	1946	50	0.79	High
TASK	1947	45	0.76	High
TASK	1948	66	0.71	High

Table A.11 Truman for Economic by Year

Trait	Year	Response Count	Mean	Classification
BACE	1945	206	0.34	Moderate
BACE	1946	260	0.38	Moderate
BACE	1947	256	0.4	High
BACE	1948	93	0.41	High
CC	1945	208	0.51	Low
CC	1946	290	0.49	Low
CC	1947	272	0.5	Low
CC	1948	90	0.51	Low

DIS	1945	115	0.02	Low
DIS	1946	169	0.03	Low
DIS	1947	149	0.07	Low
DIS	1948	57	0.03	Low
IGB	1945	74	0.08	Moderate
IGB	1946	89	0.05	Low
IGB	1947	75	0.05	Low
IGB	1948	25	0	Low
PWR	1945	196	0.32	High
PWR	1946	245	0.29	High
PWR	1947	239	0.25	Moderate
PWR	1948	87	0.19	Low
SC	1945	197	0.57	High
SC	1946	244	0.61	High
SC	1947	238	0.53	Moderate
SC	1948	91	0.5	Moderate
TASK	1945	138	0.74	High
TASK	1946	175	0.72	High
TASK	1947	180	0.82	High
TASK	1948	77	0.8	High

Table A.12 Truman for Political by Year

Trait	Year	Response Count	Mean	Classification
BACE	1945	484	0.34	Moderate
BACE	1946	414	0.37	Moderate
BACE	1947	342	0.37	Moderate
BACE	1948	399	0.37	Moderate
CC	1945	450	0.49	Low
CC	1946	336	0.46	Low
CC	1947	315	0.46	Low
CC	1948	364	0.46	Low
DIS	1945	350	0.06	Low
DIS	1946	247	0.03	Low
DIS	1947	216	0.03	Low
DIS	1948	272	0.02	Low
IGB	1945	94	0.04	Low
IGB	1946	54	0.05	Low
IGB	1947	71	0.08	Moderate
IGB	1948	63	0.06	Moderate
PWR	1945	460	0.24	Moderate
PWR	1946	386	0.22	Moderate
PWR	1947	322	0.22	Moderate
PWR	1948	381	0.2	Moderate
SC	1945	501	0.51	Moderate
SC	1946	434	0.56	High
SC	1947	362	0.53	Moderate
SC	1948	427	0.48	Moderate

TASK	1945	322	0.76	High
TASK	1946	213	0.79	High
TASK	1947	230	0.76	High
TASK	1948	257	0.76	High

Table A.13 Bush's Belief In One's Own Ability To Control Events for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	24	0.26	Low
BACE	B	45	0.3	Moderate
BACE	C	107	0.32	Moderate
BACE	D	121	0.29	Moderate
BACE	E	153	0.28	Low
BACE	F	106	0.32	Moderate
BACE	G	255	0.29	Low
BACE	H	303	0.28	Low
BACE	I	197	0.3	Moderate
BACE	J	120	0.34	Moderate
BACE	K	146	0.24	Low
BACE	L	55	0.34	Moderate
BACE	M	20	0.29	Moderate
BACE	N	55	0.41	High
BACE	O	78	0.37	Moderate
BACE	P	4	0.49	High

Table A.14 Bush's Conceptual Complexity for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	B	51	0.6	Moderate
CC	C	122	0.63	Moderate
CC	D	139	0.67	High
CC	E	155	0.64	Moderate
CC	F	106	0.67	High
CC	G	288	0.56	Moderate
CC	H	323	0.58	Moderate
CC	I	202	0.63	Moderate
CC	J	128	0.62	Moderate
CC	K	152	0.58	Moderate
CC	L	54	0.67	High
CC	M	28	0.44	Low
CC	N	64	0.58	Moderate
CC	O	80	0.57	Moderate
CC	P	5	0.46	Low

Table A.15 Bush's Distrust of Others for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	23	0.11	Moderate
DIS	B	25	0.17	High
DIS	C	90	0.18	High
DIS	D	96	0.16	High
DIS	E	118	0.07	Low
DIS	F	87	0.05	Low
DIS	G	207	0.09	Moderate
DIS	H	267	0.11	Moderate
DIS	I	177	0.13	Moderate
DIS	J	109	0.14	Moderate
DIS	K	121	0.11	Moderate
DIS	L	53	0.14	Moderate
DIS	M	19	0.07	Low
DIS	N	50	0.07	Low
DIS	O	55	0.06	Low
DIS	P	5	0.1	Moderate

Table A.16 Bush's In-group Bias for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	19	0.06	Low
IGB	B	23	0.01	Low
IGB	C	73	0.04	Low
IGB	D	97	0.03	Low
IGB	E	112	0.06	Low
IGB	F	86	0.05	Low
IGB	G	149	0.05	Low
IGB	H	173	0.07	Moderate
IGB	I	153	0.06	Low
IGB	J	94	0.07	Moderate
IGB	K	108	0.04	Low
IGB	L	50	0.1	Moderate
IGB	M	13	0.08	Moderate
IGB	O	60	0.11	High
IGB	P	4	0	Low

Table A.17 Bush's Need for Power for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	23	0.25	Moderate
PWR	B	45	0.26	Moderate
PWR	C	107	0.26	High
PWR	D	121	0.22	Moderate
PWR	E	153	0.25	Moderate
PWR	F	106	0.24	Moderate

PWR	G	255	0.23	Moderate
PWR	H	301	0.23	Moderate
PWR	I	197	0.24	Moderate
PWR	J	120	0.25	Moderate
PWR	K	146	0.18	Low
PWR	L	55	0.18	Low
PWR	M	20	0.14	Low
PWR	N	56	0.29	High
PWR	O	78	0.29	High
PWR	P	4	0.59	High

Table A.18 Bush's Self Confidence for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	19	0.52	Moderate
SC	B	46	0.5	Moderate
SC	C	106	0.55	High
SC	D	120	0.55	Moderate
SC	E	142	0.61	High
SC	F	101	0.57	High
SC	G	253	0.59	High
SC	H	304	0.55	High
SC	I	188	0.58	High
SC	J	116	0.59	High
SC	K	138	0.57	High
SC	L	53	0.53	Moderate
SC	M	18	0.51	Moderate
SC	N	52	0.55	High
SC	O	72	0.5	Moderate
SC	P	4	0.52	Moderate

Table A.19 Bush's Task Orientation for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	24	0.56	Low
TASK	B	40	0.65	Moderate
TASK	C	107	0.56	Low
TASK	D	107	0.59	Moderate
TASK	E	142	0.57	Low
TASK	F	98	0.63	Moderate
TASK	G	238	0.56	Low
TASK	H	278	0.57	Low
TASK	I	178	0.58	Low
TASK	J	121	0.63	Moderate
TASK	K	140	0.56	Low
TASK	L	55	0.48	Low
TASK	M	22	0.67	Moderate
TASK	N	57	0.63	Moderate

TASK	O	79	0.53	Low
TASK	P	5	0.73	High

Table A.20 Bush's Belief In One's Own Ability To Control Events for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	35	0.28	Low
BACE	B	19	0.38	Moderate
BACE	C	39	0.35	Moderate
BACE	D	30	0.3	Moderate
BACE	E	49	0.35	Moderate
BACE	F	78	0.35	Moderate
BACE	G	105	0.33	Moderate
BACE	H	38	0.28	Low
BACE	I	31	0.45	High
BACE	J	63	0.3	Moderate
BACE	K	54	0.29	Moderate
BACE	L	166	0.38	Moderate
BACE	M	54	0.39	High
BACE	N	58	0.33	Moderate
BACE	O	80	0.28	Low

Table A.21 Bush's Conceptual Complexity for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	35	0.64	High
CC	B	20	0.68	High
CC	C	40	0.65	High
CC	D	31	0.68	High
CC	E	52	0.69	High
CC	F	76	0.67	High
CC	G	104	0.66	High
CC	H	38	0.68	High
CC	I	34	0.68	High
CC	J	63	0.71	High
CC	K	56	0.64	High
CC	L	172	0.65	High
CC	M	57	0.65	High
CC	N	66	0.63	Moderate
CC	O	82	0.63	Moderate

Table A.22 Bush's Distrust of Others for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	25	0.12	Moderate
DIS	B	18	0.2	High
DIS	C	35	0.03	Low
DIS	D	24	0.13	Moderate

DIS	E	41	0.09	Moderate
DIS	F	67	0.08	Low
DIS	G	81	0.09	Moderate
DIS	H	35	0.18	High
DIS	I	32	0.11	Moderate
DIS	J	55	0.09	Moderate
DIS	K	44	0.07	Low
DIS	L	147	0.1	Moderate
DIS	M	52	0.06	Low
DIS	N	49	0.08	Low
DIS	O	77	0.09	Moderate

Table A.23 Bush's In-group Bias for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	24	0.12	High
IGB	B	15	0.15	High
IGB	C	29	0.07	Moderate
IGB	D	22	0.11	High
IGB	E	40	0.08	Moderate
IGB	F	60	0.05	Low
IGB	G	73	0.04	Low
IGB	H	31	0.14	High
IGB	I	32	0.05	Low
IGB	J	50	0.07	Moderate
IGB	K	37	0.05	Low
IGB	L	135	0.06	Moderate
IGB	M	35	0.06	Low
IGB	N	44	0.05	Low
IGB	O	62	0.06	Low

Table A.24 Bush's Need for Power for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	35	0.25	Moderate
PWR	B	19	0.26	Moderate
PWR	C	38	0.27	High
PWR	D	30	0.31	High
PWR	E	49	0.31	High
PWR	F	78	0.29	High
PWR	G	105	0.27	High
PWR	H	38	0.3	High
PWR	I	31	0.27	High
PWR	J	63	0.28	High
PWR	K	53	0.24	Moderate
PWR	L	165	0.3	High
PWR	M	51	0.27	High
PWR	N	58	0.32	High
PWR	O	80	0.24	Moderate

Table A.25 Bush's Self Confidence for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	35	0.56	High
SC	B	19	0.56	High
SC	C	36	0.54	Moderate
SC	D	30	0.68	High
SC	E	47	0.53	Moderate
SC	F	74	0.58	High
SC	G	103	0.53	Moderate
SC	H	33	0.64	High
SC	I	27	0.54	Moderate
SC	J	62	0.71	High
SC	K	52	0.6	High
SC	L	162	0.58	High
SC	M	52	0.56	High
SC	N	57	0.64	High
SC	O	79	0.59	High

Table A.26 Bush's Task Orientation for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	31	0.71	Moderate
TASK	B	18	0.66	Moderate
TASK	C	37	0.71	Moderate
TASK	D	26	0.63	Moderate
TASK	E	55	0.68	Moderate
TASK	F	82	0.71	Moderate
TASK	G	94	0.68	Moderate
TASK	H	34	0.75	High
TASK	I	31	0.69	Moderate
TASK	J	59	0.64	Moderate
TASK	K	51	0.61	Moderate
TASK	L	168	0.61	Moderate
TASK	M	52	0.73	High
TASK	N	57	0.69	Moderate
TASK	O	73	0.59	Moderate

Table A.27 Bush's Belief In One's Own Ability To Control Events for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	309	0.32	Moderate
BACE	B	256	0.32	Moderate
BACE	C	189	0.32	Moderate
BACE	D	280	0.33	Moderate
BACE	E	133	0.29	Moderate
BACE	F	121	0.27	Low

BACE	G	87	0.27	Low
BACE	H	68	0.31	Moderate
BACE	I	63	0.32	Moderate
BACE	J	206	0.32	Moderate
BACE	K	284	0.31	Moderate
BACE	L	202	0.34	Moderate
BACE	M	214	0.29	Low
BACE	N	274	0.29	Low
BACE	O	230	0.29	Low
BACE	P	17	0.39	High

Table A.28 Bush's Conceptual Complexity for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	329	0.6	Moderate
CC	B	280	0.67	High
CC	C	202	0.66	High
CC	D	288	0.68	High
CC	E	136	0.63	Moderate
CC	F	125	0.67	High
CC	G	80	0.6	Moderate
CC	H	77	0.66	High
CC	I	67	0.64	High
CC	J	218	0.65	High
CC	K	295	0.58	Moderate
CC	L	217	0.67	High
CC	M	219	0.6	Moderate
CC	N	272	0.63	Moderate
CC	O	236	0.68	High
CC	P	17	0.5	Low

Table A.29 Bush's Distrust of Others for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	250	0.13	Moderate
DIS	B	216	0.11	Moderate
DIS	C	162	0.11	Moderate
DIS	D	223	0.1	Moderate
DIS	E	109	0.06	Low
DIS	F	96	0.07	Low
DIS	G	67	0.06	Low
DIS	H	68	0.08	Low
DIS	I	58	0.09	Moderate
DIS	J	177	0.13	Moderate
DIS	K	250	0.09	Moderate
DIS	L	191	0.1	Moderate
DIS	M	198	0.1	Moderate
DIS	N	221	0.13	Moderate

DIS	O	218	0.1	Moderate
DIS	P	14	0.18	High

Table A.30 Bush's In-group Bias for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	188	0.06	Moderate
IGB	B	171	0.09	Moderate
IGB	C	128	0.09	Moderate
IGB	D	174	0.05	Low
IGB	E	98	0.09	Moderate
IGB	F	76	0.04	Low
IGB	G	49	0.09	Moderate
IGB	H	54	0.08	Moderate
IGB	I	50	0.05	Low
IGB	J	140	0.03	Low
IGB	K	174	0.05	Low
IGB	L	148	0.09	Moderate
IGB	M	139	0.09	Moderate
IGB	N	190	0.06	Low
IGB	O	190	0.09	Moderate
IGB	P	15	0.06	Moderate

Table A.31 Bush's Need for Power for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	308	0.24	Moderate
PWR	B	255	0.22	Moderate
PWR	C	189	0.28	High
PWR	D	272	0.24	Moderate
PWR	E	130	0.23	Moderate
PWR	F	118	0.23	Moderate
PWR	G	87	0.25	Moderate
PWR	H	68	0.25	Moderate
PWR	I	63	0.22	Moderate
PWR	J	206	0.23	Moderate
PWR	K	283	0.24	Moderate
PWR	L	201	0.28	High
PWR	M	213	0.26	High
PWR	N	272	0.26	High
PWR	O	230	0.26	High
PWR	P	17	0.22	Moderate

Table A.32 Bush's Self Confidence for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	306	0.56	High
SC	B	256	0.52	Moderate
SC	C	185	0.52	Moderate

SC	D	267	0.57	High
SC	E	125	0.55	Moderate
SC	F	119	0.52	Moderate
SC	G	84	0.53	Moderate
SC	H	69	0.48	Moderate
SC	I	62	0.54	Moderate
SC	J	203	0.48	Moderate
SC	K	274	0.6	High
SC	L	194	0.54	Moderate
SC	M	210	0.56	High
SC	N	273	0.55	High
SC	O	229	0.55	High
SC	P	17	0.31	Low

Table A.33 Bush's Task Orientation for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	299	0.62	Moderate
TASK	B	257	0.62	Moderate
TASK	C	196	0.66	Moderate
TASK	D	274	0.61	Moderate
TASK	E	120	0.7	Moderate
TASK	F	116	0.66	Moderate
TASK	G	66	0.63	Moderate
TASK	H	69	0.64	Moderate
TASK	I	62	0.59	Low
TASK	J	191	0.56	Low
TASK	K	263	0.57	Low
TASK	L	194	0.58	Low
TASK	M	196	0.62	Moderate
TASK	N	239	0.59	Moderate
TASK	O	224	0.59	Moderate
TASK	P	14	0.52	Low

Table A.34 Clinton's Belief In One's Own Ability To Control Events for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	95	0.35	Moderate
BACE	B	97	0.39	High
BACE	C	140	0.29	Moderate
BACE	D	217	0.36	Moderate
BACE	E	341	0.33	Moderate
BACE	F	336	0.37	Moderate
BACE	G	99	0.38	Moderate
BACE	H	29	0.4	High
BACE	I	5	0.43	High

BACE	J	11	0.28	Low
BACE	K	21	0.33	Moderate
BACE	L	38	0.27	Low
BACE	M	70	0.42	High
BACE	N	82	0.34	Moderate
BACE	O	24	0.44	High
BACE	P	18	0.44	High

Table A.35 Clinton's Conceptual Complexity for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	94	0.59	Moderate
CC	B	102	0.58	Moderate
CC	C	139	0.59	Moderate
CC	D	223	0.59	Moderate
CC	E	332	0.56	Moderate
CC	F	337	0.63	Moderate
CC	G	103	0.57	Moderate
CC	H	28	0.54	Low
CC	I	4	0.35	Low
CC	J	12	0.49	Low
CC	K	19	0.52	Low
CC	L	35	0.51	Low
CC	M	73	0.56	Moderate
CC	N	80	0.59	Moderate
CC	O	24	0.52	Low
CC	P	18	0.68	High

Table A.36 Clinton's Distrust of Others for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	72	0.08	Moderate
DIS	B	82	0.11	Moderate
DIS	C	121	0.09	Moderate
DIS	D	196	0.06	Low
DIS	E	289	0.06	Low
DIS	F	257	0.07	Low
DIS	G	91	0.1	Moderate
DIS	H	23	0.03	Low
DIS	I	5	0	Low
DIS	J	12	0.1	Moderate
DIS	K	16	0.08	Moderate
DIS	L	31	0.03	Low
DIS	M	57	0.13	Moderate
DIS	N	66	0.14	Moderate
DIS	O	23	0.12	Moderate
DIS	P	14	0.12	Moderate

Table A.37 Clinton's In-group Bias for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	69	0.04	Low
IGB	B	77	0.07	Moderate
IGB	C	113	0.08	Moderate
IGB	D	194	0.03	Low
IGB	E	291	0.04	Low
IGB	F	287	0.06	Moderate
IGB	G	96	0.09	Moderate
IGB	H	25	0.04	Low
IGB	I	4	0.08	Moderate
IGB	J	9	0.02	Low
IGB	K	17	0.03	Low
IGB	L	32	0.11	High
IGB	M	64	0.03	Low
IGB	N	67	0.08	Moderate
IGB	O	24	0.13	High
IGB	P	16	0.02	Low

Table A.38 Clinton's Need for Power for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	94	0.27	High
PWR	B	97	0.29	High
PWR	C	138	0.24	Moderate
PWR	D	216	0.27	High
PWR	E	336	0.26	High
PWR	F	336	0.24	Moderate
PWR	G	99	0.24	Moderate
PWR	H	29	0.2	Low
PWR	I	5	0.27	High
PWR	J	11	0.26	High
PWR	K	21	0.23	Moderate
PWR	L	38	0.26	Moderate
PWR	M	70	0.3	High
PWR	N	81	0.29	High
PWR	O	24	0.24	Moderate
PWR	P	18	0.32	High

Table A.39 Clinton's Self Confidence for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	88	0.61	High
SC	B	94	0.6	High
SC	C	130	0.61	High
SC	D	201	0.63	High
SC	E	303	0.66	High

SC	F	309	0.65	High
SC	G	90	0.58	High
SC	H	29	0.66	High
SC	I	5	0.65	High
SC	J	11	0.51	Moderate
SC	K	18	0.62	High
SC	L	38	0.62	High
SC	M	58	0.66	High
SC	N	78	0.57	High
SC	O	23	0.54	Moderate
SC	P	16	0.65	High

Table A.40 Clinton's Task Orientation for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	91	0.69	Moderate
TASK	B	94	0.65	Moderate
TASK	C	137	0.61	Moderate
TASK	D	220	0.62	Moderate
TASK	E	322	0.61	Moderate
TASK	F	326	0.65	Moderate
TASK	G	99	0.64	Moderate
TASK	H	29	0.55	Low
TASK	I	5	0.61	Moderate
TASK	J	12	0.73	High
TASK	K	20	0.61	Moderate
TASK	L	31	0.59	Moderate
TASK	M	69	0.59	Moderate
TASK	N	81	0.59	Moderate
TASK	O	25	0.61	Moderate
TASK	P	15	0.64	Moderate

Table A.41 Clinton's Belief In One's Own Ability To Control Events for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
BACE	A	53	0.37	Moderate
BACE	B	227	0.37	Moderate
BACE	C	32	0.37	Moderate
BACE	D	114	0.32	Moderate
BACE	E	231	0.36	Moderate
BACE	F	83	0.27	Low
BACE	G	11	0.33	Moderate
BACE	H	23	0.46	High
BACE	I	1	0	Low
BACE	J	3	0.36	Moderate
BACE	K	2	0.36	Moderate
BACE	M	12	0.25	Low

BACE	N	13	0.41	High
BACE	O	5	0.3	Moderate
BACE	P	2	0.54	High

Table A.42 Clinton's Conceptual Complexity for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
CC	A	53	0.68	High
CC	B	235	0.66	High
CC	C	32	0.63	Moderate
CC	D	113	0.64	Moderate
CC	E	243	0.64	Moderate
CC	F	86	0.69	High
CC	G	11	0.62	Moderate
CC	H	23	0.61	Moderate
CC	I	1	0	Low
CC	J	3	0.67	High
CC	K	2	0.4	Low
CC	M	11	0.53	Low
CC	N	13	0.74	High
CC	O	5	0.83	High
CC	P	2	0.38	Low

Table A.43 Clinton's Distrust of Others for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
DIS	A	45	0.07	Low
DIS	B	212	0.12	Moderate
DIS	C	30	0.05	Low
DIS	D	101	0.11	Moderate
DIS	E	197	0.11	Moderate
DIS	F	80	0.03	Low
DIS	G	11	0.05	Low
DIS	H	22	0.04	Low
DIS	J	3	0.12	Moderate
DIS	K	2	0	Low
DIS	M	9	0	Low
DIS	N	12	0.04	Low
DIS	O	3	0	Low
DIS	P	2	0	Low

Table A.44 Clinton's In-group Bias for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
IGB	A	48	0.03	Low
IGB	B	201	0.05	Low
IGB	C	28	0.13	High
IGB	D	96	0.09	Moderate
IGB	E	210	0.03	Low
IGB	F	85	0.04	Low
IGB	G	11	0.05	Low

IGB	H	22	0.04	Low
IGB	I	1	0	Low
IGB	J	3	0.03	Low
IGB	K	2	0.03	Low
IGB	M	10	0.22	High
IGB	N	13	0.09	Moderate
IGB	O	5	0.1	High
IGB	P	2	0	Low

Table A.45 Clinton's Need for Power for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
PWR	A	53	0.31	High
PWR	B	227	0.3	High
PWR	C	31	0.33	High
PWR	D	113	0.23	Moderate
PWR	E	231	0.23	Moderate
PWR	F	83	0.27	High
PWR	G	11	0.32	High
PWR	H	23	0.26	Moderate
PWR	I	1	0.5	High
PWR	J	3	0.2	Moderate
PWR	K	2	0.46	High
PWR	M	12	0.2	Moderate
PWR	N	13	0.13	Low
PWR	O	5	0.32	High
PWR	P	2	0.27	High

Table A.46 Clinton's Self Confidence for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
SC	A	47	0.61	High
SC	B	211	0.62	High
SC	C	29	0.64	High
SC	D	106	0.63	High
SC	E	216	0.56	High
SC	F	77	0.58	High
SC	G	11	0.61	High
SC	H	22	0.62	High
SC	I	1	1	High
SC	J	3	0.65	High
SC	K	2	0.92	High
SC	M	12	0.77	High
SC	N	9	0.83	High
SC	O	5	0.51	Moderate
SC	P	2	0.34	Low

Table A.47 Clinton's Task Orientation for Economic Responses by Quarter

Trait	Quarter	Count	Mean	Classification
TASK	A	53	0.66	Moderate

TASK	B	225	0.67	Moderate
TASK	C	33	0.66	Moderate
TASK	D	110	0.66	Moderate
TASK	E	237	0.66	Moderate
TASK	F	83	0.62	Moderate
TASK	G	11	0.62	Moderate
TASK	H	23	0.77	High
TASK	I	1	0	Low
TASK	J	3	0.73	High
TASK	K	2	0.59	Low
TASK	M	12	0.49	Low
TASK	N	12	0.71	High
TASK	O	5	0.87	High
TASK	P	2	0.58	Low

Table A.48 Clinton’s Belief In One’s Own Ability To Control Events for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	37	0.37	Moderate
BACE	B	94	0.38	Moderate
BACE	C	94	0.36	Moderate
BACE	D	131	0.36	Moderate
BACE	E	251	0.37	Moderate
BACE	F	186	0.39	Moderate
BACE	G	63	0.35	Moderate
BACE	H	42	0.36	Moderate
BACE	I	13	0.23	Low
BACE	J	12	0.24	Low
BACE	K	39	0.39	Moderate
BACE	L	13	0.42	High
BACE	M	37	0.3	Moderate
BACE	N	95	0.32	Moderate
BACE	O	21	0.52	High
BACE	P	7	0.19	Low

Table A.49 Clinton’s Conceptual Complexity for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	36	0.55	Moderate
CC	B	99	0.56	Moderate
CC	C	96	0.62	Moderate
CC	D	128	0.59	Moderate
CC	E	255	0.53	Low
CC	F	185	0.58	Moderate
CC	G	66	0.57	Moderate
CC	H	43	0.62	Moderate

CC	I	11	0.53	Low
CC	J	11	0.61	Moderate
CC	K	41	0.63	Moderate
CC	L	9	0.62	Moderate
CC	M	40	0.54	Low
CC	N	97	0.58	Moderate
CC	O	21	0.45	Low
CC	P	9	0.3	Low

Table A.50 Clinton’s Distrust of Others for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	30	0.07	Low
DIS	B	86	0.1	Moderate
DIS	C	86	0.12	Moderate
DIS	D	121	0.11	Moderate
DIS	E	226	0.12	Moderate
DIS	F	145	0.08	Low
DIS	G	56	0.08	Moderate
DIS	H	42	0.07	Low
DIS	I	10	0.08	Moderate
DIS	J	9	0.12	Moderate
DIS	K	33	0.11	Moderate
DIS	L	8	0.04	Low
DIS	M	33	0.09	Moderate
DIS	N	88	0.07	Low
DIS	O	14	0.13	Moderate
DIS	P	5	0.27	High

Table A.51 Clinton’s In-group Bias for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	16	0.03	Low
IGB	B	63	0.09	Moderate
IGB	C	75	0.05	Low
IGB	D	96	0.12	High
IGB	E	190	0.08	Moderate
IGB	F	147	0.06	Low
IGB	G	57	0.03	Low
IGB	H	38	0.05	Low
IGB	I	9	0.03	Low
IGB	J	7	0.02	Low
IGB	K	32	0.04	Low
IGB	L	9	0	Low
IGB	M	25	0.03	Low
IGB	N	67	0.04	Low
IGB	O	17	0.15	High
IGB	P	6	0.08	Moderate

Table A.52 Clinton's Need for Power for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	36	0.29	High
PWR	B	93	0.33	High
PWR	C	92	0.27	High
PWR	D	130	0.26	Moderate
PWR	E	250	0.25	Moderate
PWR	F	185	0.27	High
PWR	G	63	0.26	Moderate
PWR	H	42	0.23	Moderate
PWR	I	13	0.23	Moderate
PWR	J	12	0.21	Moderate
PWR	K	38	0.29	High
PWR	L	13	0.2	Moderate
PWR	M	36	0.18	Low
PWR	N	94	0.25	Moderate
PWR	O	21	0.32	High
PWR	P	6	0.19	Low

Table A.53 Clinton's Self Confidence for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	37	0.63	High
SC	B	88	0.54	Moderate
SC	C	93	0.56	High
SC	D	130	0.56	High
SC	E	235	0.51	Moderate
SC	F	187	0.6	High
SC	G	62	0.61	High
SC	H	43	0.56	High
SC	I	13	0.6	High
SC	J	12	0.49	Moderate
SC	K	40	0.5	Moderate
SC	L	12	0.7	High
SC	M	34	0.56	High
SC	N	92	0.57	High
SC	O	17	0.62	High
SC	P	6	0.79	High

Table A.54 Clinton's Task Orientation for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	32	0.57	Low
TASK	B	84	0.65	Moderate
TASK	C	93	0.62	Moderate
TASK	D	124	0.61	Moderate

TASK	E	236	0.63	Moderate
TASK	F	185	0.61	Moderate
TASK	G	66	0.63	Moderate
TASK	H	42	0.6	Moderate
TASK	I	9	0.55	Low
TASK	J	9	0.68	Moderate
TASK	K	37	0.56	Low
TASK	L	12	0.76	High
TASK	M	37	0.64	Moderate
TASK	N	90	0.61	Moderate
TASK	O	20	0.62	Moderate
TASK	P	7	0.7	Moderate

Table A.55 Johnson's Belief In One's Own Ability To Control Events for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	12	0.24	Low
BACE	B	54	0.36	Moderate
BACE	C	24	0.36	Moderate
BACE	D	20	0.38	Moderate
BACE	E	12	0.38	Moderate
BACE	F	24	0.31	Moderate
BACE	G	43	0.3	Moderate
BACE	I	26	0.22	Low
BACE	J	22	0.39	Moderate
BACE	K	38	0.3	Moderate
BACE	L	41	0.32	Moderate
BACE	M	34	0.35	Moderate
BACE	N	63	0.34	Moderate
BACE	O	35	0.36	Moderate
BACE	P	37	0.39	High

Table A.56 Johnson's Conceptual Complexity for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	12	0.49	Low
CC	B	54	0.57	Moderate
CC	C	24	0.49	Low
CC	D	20	0.53	Low
CC	E	12	0.56	Moderate
CC	F	23	0.6	Moderate
CC	G	41	0.59	Moderate
CC	I	25	0.45	Low
CC	J	23	0.61	Moderate
CC	K	36	0.59	Moderate
CC	L	41	0.55	Moderate
CC	M	32	0.55	Moderate

CC	N	56	0.59	Moderate
CC	O	31	0.62	Moderate
CC	P	38	0.55	Moderate

Table A.57 Johnson's Distrust of Others for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	9	0.05	Low
DIS	B	42	0.09	Moderate
DIS	C	14	0.04	Low
DIS	D	12	0.06	Low
DIS	E	7	0.03	Low
DIS	F	19	0.06	Low
DIS	G	26	0.06	Low
DIS	I	15	0	Low
DIS	J	17	0.04	Low
DIS	K	24	0.2	High
DIS	L	28	0.21	High
DIS	M	21	0.11	Moderate
DIS	N	38	0.12	Moderate
DIS	O	22	0.07	Low
DIS	P	26	0.12	Moderate

Table A.58 Johnson's In-group Bias for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	11	0.08	Moderate
IGB	B	40	0.13	High
IGB	C	17	0.02	Low
IGB	D	11	0.08	Moderate
IGB	E	10	0.15	High
IGB	F	21	0.08	Moderate
IGB	G	38	0.09	Moderate
IGB	I	15	0.05	Low
IGB	J	18	0.02	Low
IGB	K	27	0.02	Low
IGB	L	33	0.05	Low
IGB	M	26	0.07	Moderate
IGB	N	46	0.05	Low
IGB	O	25	0.05	Low
IGB	P	31	0.03	Low

Table A.59 Johnson's Need for Power for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	12	0.16	Low
PWR	B	53	0.21	Moderate
PWR	C	24	0.18	Low

PWR	D	20	0.16	Low
PWR	E	12	0.27	High
PWR	F	24	0.2	Moderate
PWR	G	43	0.22	Moderate
PWR	I	25	0.2	Low
PWR	J	22	0.32	High
PWR	K	38	0.23	Moderate
PWR	L	41	0.22	Moderate
PWR	M	34	0.17	Low
PWR	N	63	0.23	Moderate
PWR	O	34	0.24	Moderate
PWR	P	37	0.26	High

Table A.60 Johnson's Self Confidence for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	11	0.49	Moderate
SC	B	48	0.54	Moderate
SC	C	21	0.64	High
SC	D	20	0.61	High
SC	E	11	0.54	Moderate
SC	F	21	0.7	High
SC	G	39	0.61	High
SC	I	25	0.63	High
SC	J	22	0.51	Moderate
SC	K	34	0.66	High
SC	L	37	0.6	High
SC	M	33	0.49	Moderate
SC	N	57	0.5	Moderate
SC	O	29	0.6	High
SC	P	36	0.63	High

Table A.61 Johnson's Task Orientation for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	12	0.63	Moderate
TASK	B	51	0.69	Moderate
TASK	C	22	0.68	Moderate
TASK	D	21	0.76	High
TASK	E	12	0.57	Low
TASK	F	23	0.61	Moderate
TASK	G	40	0.64	Moderate
TASK	I	21	0.49	Low
TASK	J	20	0.64	Moderate
TASK	K	35	0.74	High
TASK	L	37	0.6	Moderate
TASK	M	29	0.47	Low
TASK	N	46	0.76	High

TASK	O	28	0.67	Moderate
TASK	P	35	0.57	Low

Table A.62 Johnson's Belief In One's Own Ability To Control Events for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	15	0.29	Low
BACE	B	58	0.36	Moderate
BACE	C	11	0.43	High
BACE	D	6	0.55	High
BACE	E	5	0.36	Moderate
BACE	F	11	0.32	Moderate
BACE	G	13	0.52	High
BACE	I	17	0.44	High
BACE	J	19	0.42	High
BACE	K	25	0.34	Moderate
BACE	L	75	0.36	Moderate
BACE	M	34	0.42	High
BACE	N	15	0.34	Moderate
BACE	O	22	0.29	Moderate
BACE	P	33	0.39	Moderate

Table A.63 Johnson's Conceptual Complexity for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	15	0.62	Moderate
CC	B	56	0.64	High
CC	C	10	0.56	Moderate
CC	D	6	0.68	High
CC	E	5	0.71	High
CC	F	10	0.62	Moderate
CC	G	13	0.6	Moderate
CC	I	15	0.58	Moderate
CC	J	18	0.64	High
CC	K	23	0.67	High
CC	L	75	0.65	High
CC	M	38	0.72	High
CC	N	15	0.68	High
CC	O	20	0.74	High
CC	P	31	0.59	Moderate

Table A.64 Johnson's Distrust of Others for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	14	0	Low
DIS	B	38	0.03	Low

DIS	C	7	0.29	High
DIS	D	3	0.1	Moderate
DIS	E	3	0	Low
DIS	F	9	0.03	Low
DIS	G	9	0.09	Moderate
DIS	I	12	0.01	Low
DIS	J	11	0.09	Moderate
DIS	K	20	0.06	Low
DIS	L	44	0.07	Low
DIS	M	22	0.03	Low
DIS	N	13	0.03	Low
DIS	O	16	0.08	Low
DIS	P	25	0.13	Moderate

Table A.65 Johnson's In-group Bias for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	12	0.07	Moderate
IGB	B	53	0.05	Low
IGB	C	7	0.17	High
IGB	D	3	0	Low
IGB	E	4	0.17	High
IGB	F	8	0.02	Low
IGB	G	11	0.05	Low
IGB	I	14	0.02	Low
IGB	J	14	0.01	Low
IGB	K	23	0.03	Low
IGB	L	61	0.04	Low
IGB	M	29	0.01	Low
IGB	N	12	0.03	Low
IGB	O	18	0.08	Moderate
IGB	P	28	0.03	Low

Table A.66 Johnson's Need for Power for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	15	0.23	Moderate
PWR	B	58	0.31	High
PWR	C	11	0.23	Moderate
PWR	D	6	0.19	Low
PWR	E	5	0.25	Moderate
PWR	F	11	0.19	Low
PWR	G	13	0.32	High
PWR	I	17	0.3	High
PWR	J	19	0.2	Moderate
PWR	K	25	0.24	Moderate
PWR	L	74	0.23	Moderate
PWR	M	34	0.2	Moderate

PWR	N	15	0.24	Moderate
PWR	O	22	0.23	Moderate
PWR	P	33	0.15	Low

Table A.67 Johnson's Self Confidence for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	14	0.62	High
SC	B	43	0.51	Moderate
SC	C	9	0.45	Moderate
SC	D	6	0.67	High
SC	E	5	0.67	High
SC	F	11	0.57	High
SC	G	13	0.4	Moderate
SC	I	16	0.55	Moderate
SC	J	18	0.48	Moderate
SC	K	20	0.62	High
SC	L	61	0.57	High
SC	M	25	0.69	High
SC	N	14	0.59	High
SC	O	20	0.57	High
SC	P	29	0.59	High

Table A.68 Johnson's Task Orientation for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	16	0.79	High
TASK	B	52	0.67	Moderate
TASK	C	9	0.62	Moderate
TASK	D	6	0.52	Low
TASK	E	5	0.59	Moderate
TASK	F	12	0.68	Moderate
TASK	G	12	0.73	High
TASK	I	15	0.75	High
TASK	J	14	0.66	Moderate
TASK	K	24	0.72	High
TASK	L	61	0.71	Moderate
TASK	M	29	0.85	High
TASK	N	15	0.52	Low
TASK	O	18	0.78	High
TASK	P	32	0.62	Moderate

Table A.69 Johnson's Belief In One's Own Ability To Control Events for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	34	0.43	High

BACE	B	169	0.31	Moderate
BACE	C	88	0.33	Moderate
BACE	D	49	0.38	Moderate
BACE	E	19	0.43	High
BACE	F	29	0.28	Low
BACE	G	32	0.32	Moderate
BACE	I	14	0.36	Moderate
BACE	J	25	0.29	Low
BACE	K	59	0.34	Moderate
BACE	L	83	0.35	Moderate
BACE	M	30	0.32	Moderate
BACE	N	19	0.33	Moderate
BACE	O	38	0.26	Low
BACE	P	44	0.32	Moderate

Table A.70 Johnson’s Conceptual Complexity for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	28	0.56	Moderate
CC	B	160	0.54	Moderate
CC	C	76	0.52	Low
CC	D	39	0.5	Low
CC	E	16	0.5	Low
CC	F	34	0.57	Moderate
CC	G	32	0.58	Moderate
CC	I	13	0.51	Low
CC	J	25	0.64	Moderate
CC	K	55	0.58	Moderate
CC	L	79	0.58	Moderate
CC	M	26	0.65	High
CC	N	20	0.57	Moderate
CC	O	33	0.49	Low
CC	P	39	0.59	Moderate

Table A.71 Johnson’s Distrust of Others for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	22	0.03	Low
DIS	B	115	0.04	Low
DIS	C	58	0.1	Moderate
DIS	D	30	0.04	Low
DIS	E	9	0	Low
DIS	F	25	0.01	Low
DIS	G	27	0.04	Low
DIS	I	12	0	Low
DIS	J	21	0.02	Low
DIS	K	46	0.04	Low
DIS	L	67	0.04	Low

DIS	M	15	0.01	Low
DIS	N	18	0.06	Low
DIS	O	26	0	Low
DIS	P	35	0.07	Low

Table A.72 Johnson's In-group Bias for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	15	0.14	High
IGB	B	92	0.05	Low
IGB	C	35	0.09	Moderate
IGB	D	31	0.01	Low
IGB	E	7	0.02	Low
IGB	F	21	0.06	Low
IGB	G	23	0.08	Moderate
IGB	I	9	0.11	High
IGB	J	16	0.07	Moderate
IGB	K	42	0.07	Moderate
IGB	L	61	0.03	Low
IGB	M	15	0.04	Low
IGB	N	10	0.03	Low
IGB	O	24	0.06	Moderate
IGB	P	23	0.06	Moderate

Table A.73 Johnson's Need for Power for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	34	0.23	Moderate
PWR	B	168	0.21	Moderate
PWR	C	87	0.21	Moderate
PWR	D	48	0.19	Low
PWR	E	17	0.21	Moderate
PWR	F	29	0.29	High
PWR	G	32	0.23	Moderate
PWR	I	14	0.12	Low
PWR	J	25	0.16	Low
PWR	K	59	0.21	Moderate
PWR	L	83	0.27	High
PWR	M	30	0.23	Moderate
PWR	N	19	0.2	Low
PWR	O	37	0.19	Low
PWR	P	44	0.22	Moderate

Table A.74 Johnson's Self Confidence for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	32	0.47	Moderate
SC	B	158	0.59	High

SC	C	84	0.57	High
SC	D	45	0.49	Moderate
SC	E	19	0.51	Moderate
SC	F	27	0.54	Moderate
SC	G	30	0.53	Moderate
SC	I	11	0.64	High
SC	J	25	0.53	Moderate
SC	K	57	0.5	Moderate
SC	L	76	0.6	High
SC	M	31	0.58	High
SC	N	19	0.51	Moderate
SC	O	34	0.53	Moderate
SC	P	44	0.52	Moderate

Table A.75 Johnson’s Task Orientation for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	28	0.71	High
TASK	B	145	0.63	Moderate
TASK	C	66	0.6	Moderate
TASK	D	34	0.55	Low
TASK	E	13	0.48	Low
TASK	F	32	0.54	Low
TASK	G	30	0.6	Moderate
TASK	I	11	0.78	High
TASK	J	22	0.65	Moderate
TASK	K	53	0.69	Moderate
TASK	L	71	0.66	Moderate
TASK	M	23	0.69	Moderate
TASK	N	17	0.67	Moderate
TASK	O	29	0.66	Moderate
TASK	P	33	0.6	Moderate

Table A.76 Truman’s Belief In One’s Own Ability To Control Events for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	11	0.5	High
BACE	B	23	0.27	Low
BACE	C	52	0.33	Moderate
BACE	D	89	0.3	Moderate
BACE	E	14	0.36	Moderate
BACE	F	16	0.31	Moderate
BACE	G	20	0.18	Low
BACE	H	20	0.29	Low
BACE	I	8	0.12	Low
BACE	J	12	0.4	High

BACE	K	14	0.59	High
BACE	L	34	0.36	Moderate
BACE	M	26	0.27	Low
BACE	N	35	0.45	High
BACE	O	21	0.08	Low
BACE	P	27	0.44	High

Table A.77 Truman's Conceptual Complexity for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	8	0.33	Low
CC	B	19	0.49	Low
CC	C	51	0.44	Low
CC	D	72	0.43	Low
CC	E	9	0.59	Moderate
CC	F	9	0.35	Low
CC	G	15	0.48	Low
CC	H	15	0.49	Low
CC	I	6	0.08	Low
CC	J	13	0.39	Low
CC	K	12	0.44	Low
CC	L	32	0.51	Low
CC	M	22	0.47	Low
CC	N	32	0.46	Low
CC	O	21	0.39	Low
CC	P	23	0.55	Moderate

Table A.78 Truman's Distrust of Others for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	5	0	Low
DIS	B	16	0	Low
DIS	C	43	0.01	Low
DIS	D	54	0.02	Low
DIS	E	9	0	Low
DIS	F	7	0	Low
DIS	G	16	0.01	Low
DIS	H	8	0.12	Moderate
DIS	I	4	0	Low
DIS	J	10	0	Low
DIS	K	6	0	Low
DIS	L	24	0.02	Low
DIS	M	17	0.06	Low
DIS	N	18	0.07	Low
DIS	O	13	0	Low
DIS	P	18	0.01	Low

Table A.79 Truman's In-group Bias for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	2	0	Low
IGB	B	3	0	Low
IGB	C	27	0.04	Low
IGB	D	23	0.04	Low
IGB	E	6	0	Low
IGB	F	2	0	Low
IGB	G	5	0.07	Moderate
IGB	H	1	0	Low
IGB	I	1	0	Low
IGB	J	6	0.33	High
IGB	K	5	0	Low
IGB	L	15	0.06	Low
IGB	M	8	0.16	High
IGB	N	10	0.12	High
IGB	O	2	0	Low
IGB	P	6	0	Low

Table A.80 Truman's Need for Power for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	9	0.16	Low
PWR	B	22	0.12	Low
PWR	C	49	0.2	Low
PWR	D	86	0.24	Moderate
PWR	E	12	0.17	Low
PWR	F	14	0.2	Moderate
PWR	G	19	0.18	Low
PWR	H	18	0.14	Low
PWR	I	8	0.1	Low
PWR	J	11	0.38	High
PWR	K	11	0.17	Low
PWR	L	30	0.23	Moderate
PWR	M	24	0.17	Low
PWR	N	31	0.37	High
PWR	O	20	0.04	Low
PWR	P	25	0.19	Low

Table A.81 Truman's Self Confidence for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	11	0.34	Low
SC	B	22	0.67	High
SC	C	41	0.61	High
SC	D	87	0.54	Moderate
SC	E	19	0.55	High
SC	F	16	0.7	High

SC	G	18	0.63	High
SC	H	21	0.47	Moderate
SC	I	10	0.46	Moderate
SC	J	12	0.5	Moderate
SC	K	13	0.7	High
SC	L	26	0.5	Moderate
SC	M	27	0.74	High
SC	N	30	0.4	Moderate
SC	O	22	0.66	High
SC	P	27	0.6	High

Table A.82 Truman's Task Orientation for Conflict Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	8	0.56	Low
TASK	B	15	0.87	High
TASK	C	36	0.7	Moderate
TASK	D	62	0.75	High
TASK	E	16	0.81	High
TASK	F	9	0.78	High
TASK	G	9	0.83	High
TASK	H	16	0.75	High
TASK	I	4	0.5	Low
TASK	J	11	0.8	High
TASK	K	4	0.75	High
TASK	L	26	0.78	High
TASK	M	15	0.65	Moderate
TASK	N	23	0.68	Moderate
TASK	O	13	0.85	High
TASK	P	15	0.69	Moderate

Table A.83 Truman's Belief In One's Own Ability To Control Events for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	11	0.38	Moderate
BACE	B	4	0.42	High
BACE	C	90	0.31	Moderate
BACE	D	101	0.37	Moderate
BACE	E	48	0.45	High
BACE	F	65	0.45	High
BACE	G	60	0.4	High
BACE	H	87	0.27	Low
BACE	I	54	0.35	Moderate
BACE	J	87	0.42	High
BACE	K	87	0.42	High
BACE	L	28	0.41	High
BACE	M	13	0.6	High

BACE	N	18	0.38	Moderate
BACE	O	29	0.43	High
BACE	P	33	0.32	Moderate

Table A.84 Truman's Conceptual Complexity for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	10	0.57	Moderate
CC	B	8	0.57	Moderate
CC	C	94	0.51	Low
CC	D	96	0.5	Low
CC	E	44	0.49	Low
CC	F	81	0.54	Moderate
CC	G	86	0.48	Low
CC	H	79	0.45	Low
CC	I	41	0.5	Low
CC	J	108	0.53	Low
CC	K	92	0.47	Low
CC	L	31	0.49	Low
CC	M	19	0.48	Low
CC	N	13	0.61	Moderate
CC	O	27	0.53	Low
CC	P	31	0.47	Low

Table A.85 Truman's Distrust of Others for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	8	0	Low
DIS	B	4	0.08	Moderate
DIS	C	54	0.04	Low
DIS	D	49	0.01	Low
DIS	E	37	0.03	Low
DIS	F	38	0.04	Low
DIS	G	45	0	Low
DIS	H	49	0.05	Low
DIS	I	24	0.08	Low
DIS	J	49	0.07	Low
DIS	K	58	0.06	Low
DIS	L	18	0.06	Low
DIS	M	9	0	Low
DIS	N	12	0.03	Low
DIS	O	22	0.06	Low
DIS	P	14	0.02	Low

Table A.86 Truman's In-group Bias for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	7	0.14	High

IGB	B	1	0	Low
IGB	C	29	0.12	High
IGB	D	37	0.05	Low
IGB	E	14	0.05	Low
IGB	F	28	0.07	Moderate
IGB	G	25	0.03	Low
IGB	H	22	0.03	Low
IGB	I	8	0.04	Low
IGB	J	36	0.01	Low
IGB	K	24	0.08	Moderate
IGB	L	7	0.14	High
IGB	M	4	0	Low
IGB	N	6	0	Low
IGB	O	6	0	Low
IGB	P	9	0	Low

Table A.87 Truman's Need for Power for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	10	0.37	High
PWR	B	3	0.56	High
PWR	C	86	0.34	High
PWR	D	97	0.28	High
PWR	E	43	0.33	High
PWR	F	61	0.31	High
PWR	G	56	0.24	Moderate
PWR	H	85	0.27	High
PWR	I	52	0.24	Moderate
PWR	J	77	0.22	Moderate
PWR	K	83	0.23	Moderate
PWR	L	27	0.4	High
PWR	M	12	0.17	Low
PWR	N	17	0.4	High
PWR	O	27	0.12	Low
PWR	P	31	0.15	Low

Table A.88 Truman's Self Confidence for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	11	0.41	Moderate
SC	B	4	0.5	Moderate
SC	C	84	0.56	High
SC	D	98	0.61	High
SC	E	49	0.61	High
SC	F	55	0.6	High
SC	G	55	0.57	High
SC	H	85	0.64	High
SC	I	56	0.62	High

SC	J	76	0.47	Moderate
SC	K	78	0.5	Moderate
SC	L	28	0.57	High
SC	M	11	0.33	Low
SC	N	18	0.61	High
SC	O	30	0.36	Low
SC	P	32	0.63	High

Table A.89 Truman's Task Orientation for Economic Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
TASK	A	9	0.9	High
TASK	B	6	0.83	High
TASK	C	54	0.69	Moderate
TASK	D	69	0.75	High
TASK	E	25	0.71	Moderate
TASK	F	49	0.77	High
TASK	G	50	0.7	Moderate
TASK	H	51	0.71	Moderate
TASK	I	25	0.84	High
TASK	J	71	0.82	High
TASK	K	64	0.79	High
TASK	L	20	0.87	High
TASK	M	14	0.82	High
TASK	N	14	0.43	Low
TASK	O	27	0.89	High
TASK	P	22	0.91	High

Table A.90 Truman's Belief In One's Own Ability To Control Events for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
BACE	A	128	0.42	High
BACE	B	124	0.29	Moderate
BACE	C	88	0.37	Moderate
BACE	D	144	0.3	Moderate
BACE	E	165	0.39	High
BACE	F	97	0.35	Moderate
BACE	G	60	0.36	Moderate
BACE	H	92	0.35	Moderate
BACE	I	91	0.39	Moderate
BACE	J	77	0.36	Moderate
BACE	K	78	0.31	Moderate
BACE	L	96	0.42	High
BACE	M	76	0.49	High
BACE	N	76	0.38	Moderate
BACE	O	94	0.43	High
BACE	P	153	0.27	Low

Table A.91 Truman's Conceptual Complexity for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
CC	A	118	0.48	Low
CC	B	126	0.52	Low
CC	C	78	0.55	Moderate
CC	D	128	0.43	Low
CC	E	129	0.49	Low
CC	F	74	0.47	Low
CC	G	53	0.42	Low
CC	H	80	0.42	Low
CC	I	79	0.41	Low
CC	J	76	0.38	Low
CC	K	71	0.53	Low
CC	L	89	0.51	Low
CC	M	65	0.48	Low
CC	N	57	0.51	Low
CC	O	101	0.47	Low
CC	P	141	0.42	Low

Table A.92 Truman's Distrust of Others for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
DIS	A	83	0.09	Moderate
DIS	B	97	0.06	Low
DIS	C	65	0.05	Low
DIS	D	105	0.04	Low
DIS	E	98	0.03	Low
DIS	F	57	0.02	Low
DIS	G	36	0.01	Low
DIS	H	56	0.06	Low
DIS	I	57	0.03	Low
DIS	J	50	0.01	Low
DIS	K	45	0.02	Low
DIS	L	64	0.04	Low
DIS	M	49	0	Low
DIS	N	40	0.03	Low
DIS	O	74	0.01	Low
DIS	P	109	0.03	Low

Table A.93 Truman's In-group Bias for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
IGB	A	28	0	Low
IGB	B	30	0.07	Moderate
IGB	C	20	0	Low
IGB	D	16	0.08	Moderate
IGB	E	20	0.08	Moderate

IGB	F	11	0.09	Moderate
IGB	G	10	0.03	Low
IGB	H	13	0	Low
IGB	I	22	0.12	High
IGB	J	28	0.06	Moderate
IGB	K	8	0	Low
IGB	L	13	0.12	High
IGB	M	5	0.1	High
IGB	N	14	0.07	Moderate
IGB	O	14	0.07	Moderate
IGB	P	30	0.05	Low

Table A.94 Truman's Need for Power for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
PWR	A	123	0.31	High
PWR	B	118	0.2	Low
PWR	C	81	0.27	High
PWR	D	138	0.18	Low
PWR	E	152	0.21	Moderate
PWR	F	90	0.28	High
PWR	G	57	0.26	Moderate
PWR	H	87	0.17	Low
PWR	I	89	0.24	Moderate
PWR	J	75	0.26	High
PWR	K	74	0.2	Moderate
PWR	L	84	0.16	Low
PWR	M	68	0.28	High
PWR	N	73	0.31	High
PWR	O	92	0.14	Low
PWR	P	148	0.16	Low

Table A.95 Truman's Self Confidence for Political Responses by Quarter

Trait	Quarter	Response Count	Mean	Classification
SC	A	130	0.5	Moderate
SC	B	128	0.59	High
SC	C	84	0.48	Moderate
SC	D	159	0.48	Moderate
SC	E	166	0.55	Moderate
SC	F	102	0.56	High
SC	G	60	0.59	High
SC	H	106	0.56	High
SC	I	93	0.56	High
SC	J	80	0.56	High
SC	K	82	0.57	High
SC	L	107	0.46	Moderate
SC	M	91	0.41	Moderate

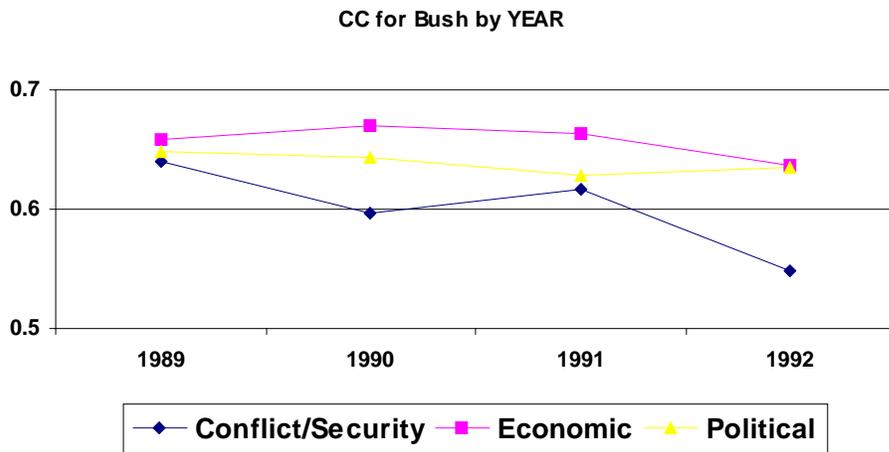
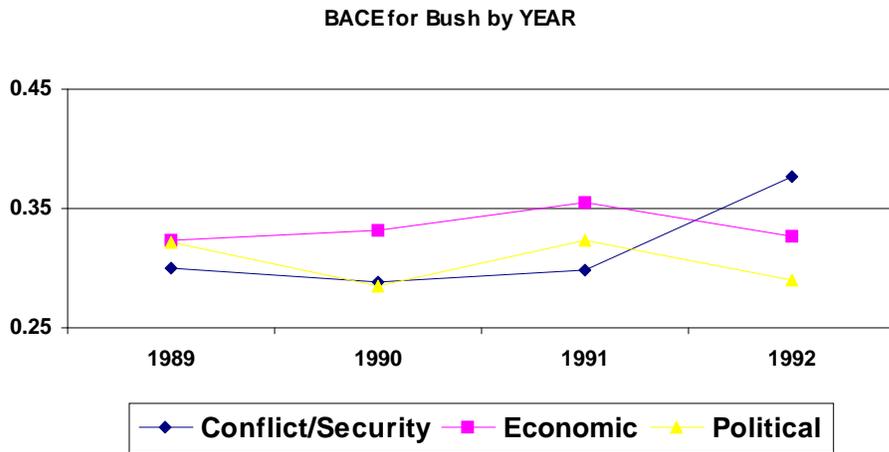
SC	N	75	0.39	Moderate
SC	O	103	0.55	High
SC	P	158	0.51	Moderate

Table A.96 Truman's Task Orientation for Political Responses by Quarter

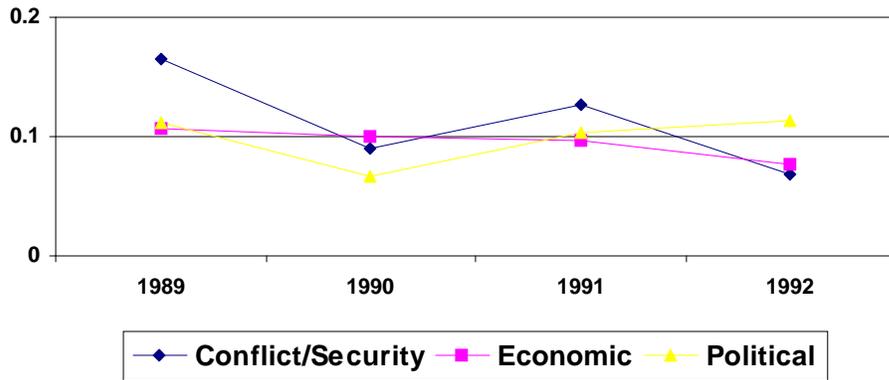
Trait	Quarter	Response Count	Mean	Classification
TASK	A	76	0.61	Moderate
TASK	B	97	0.87	High
TASK	C	50	0.85	High
TASK	D	99	0.71	Moderate
TASK	E	82	0.77	High
TASK	F	50	0.74	High
TASK	G	42	0.9	High
TASK	H	39	0.77	High
TASK	I	53	0.64	Moderate
TASK	J	63	0.85	High
TASK	K	51	0.74	High
TASK	L	63	0.78	High
TASK	M	40	0.85	High
TASK	N	39	0.72	High
TASK	O	75	0.69	Moderate
TASK	P	103	0.8	High

APPENDIX B

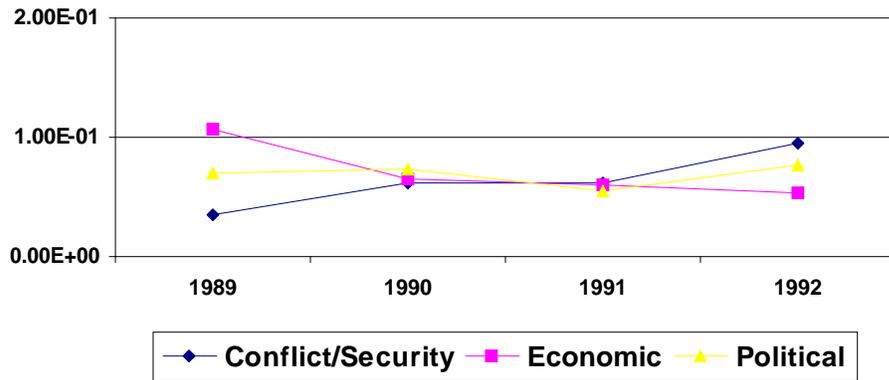
Graphical Representation of Trait Score Variation Across Time



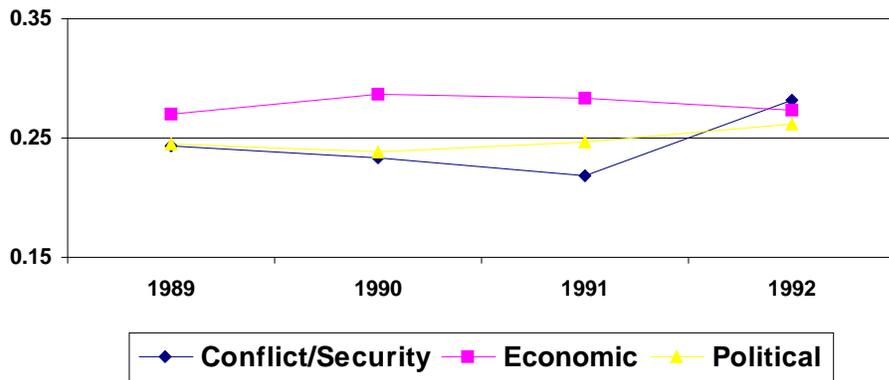
DIS for Bush by YEAR



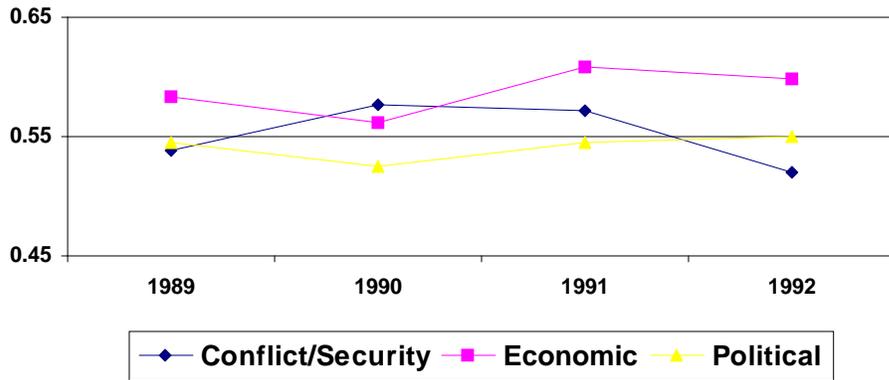
IGB for Bush by YEAR



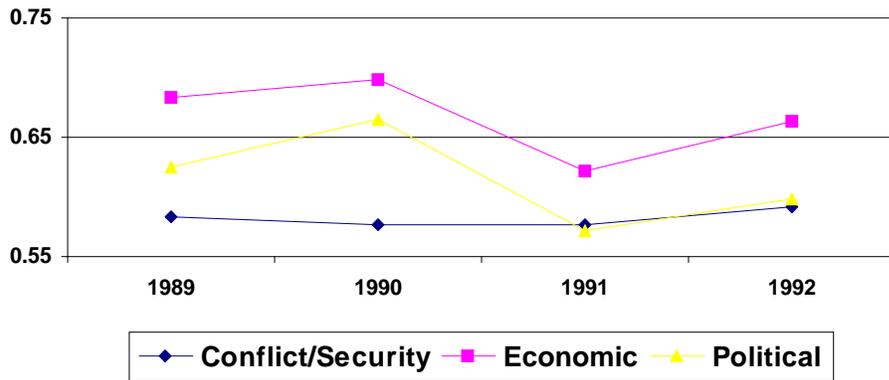
PWR for Bush by YEAR



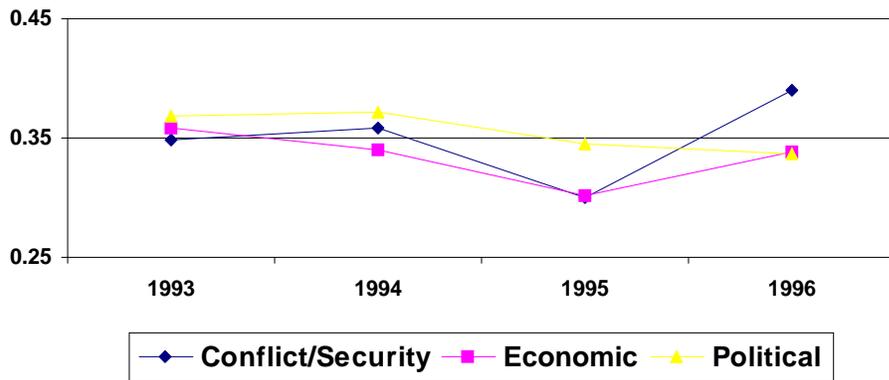
SC for Bush by YEAR



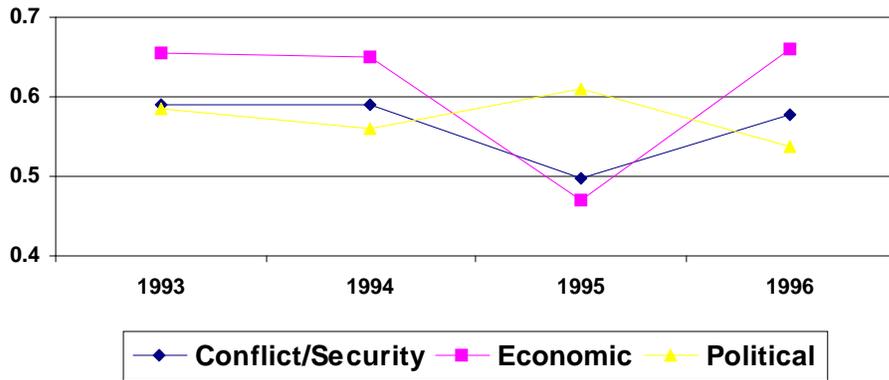
TASK for Bush by YEAR



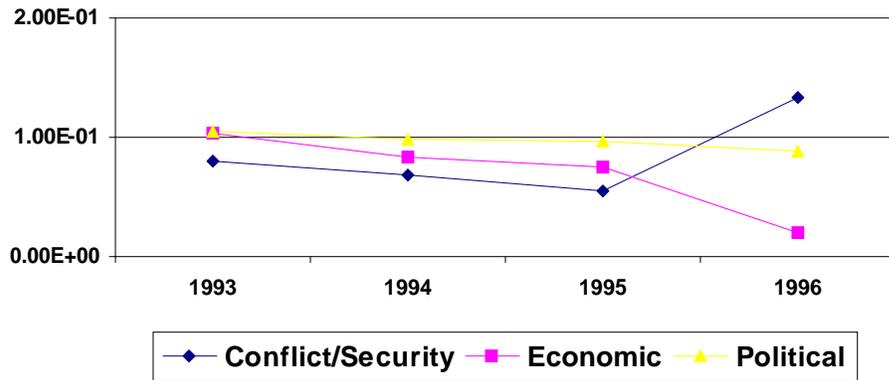
BACE for Clinton by YEAR



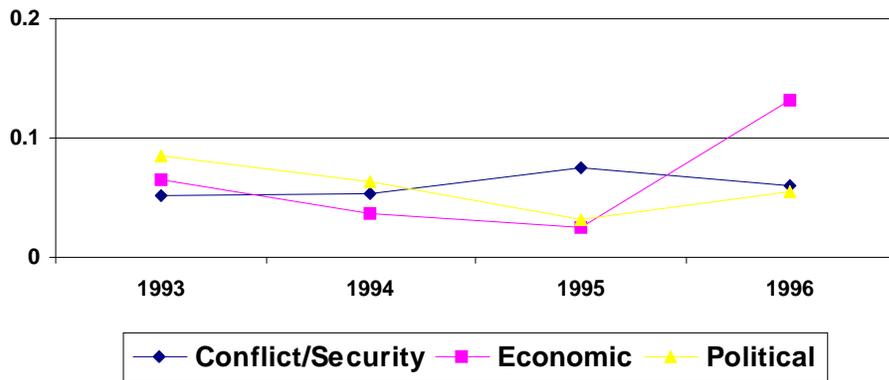
CC for Clinton by YEAR



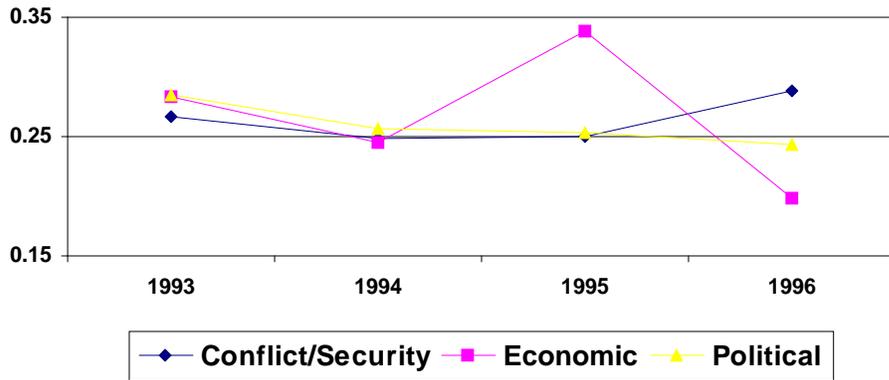
DIS for Clinton by YEAR



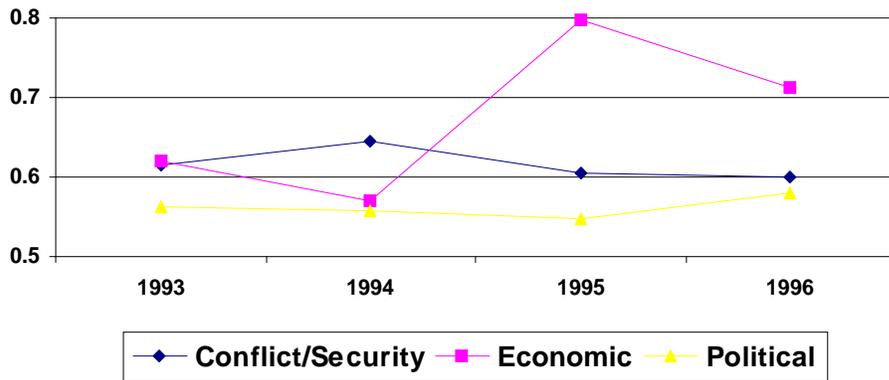
IGB for Clinton by YEAR



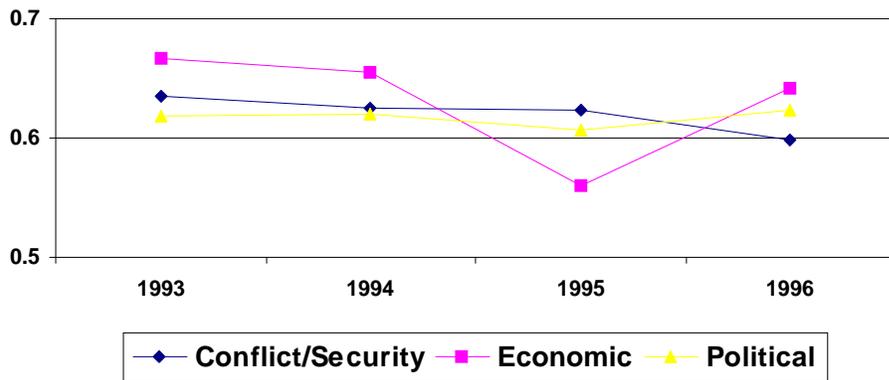
PWR for Clinton by YEAR



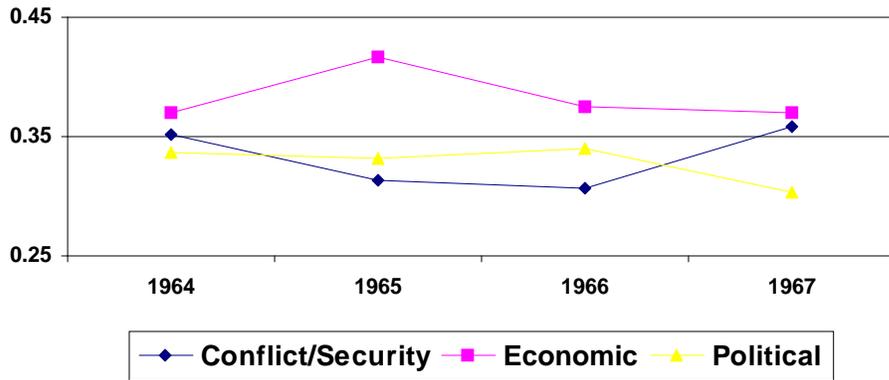
SC for Clinton by YEAR



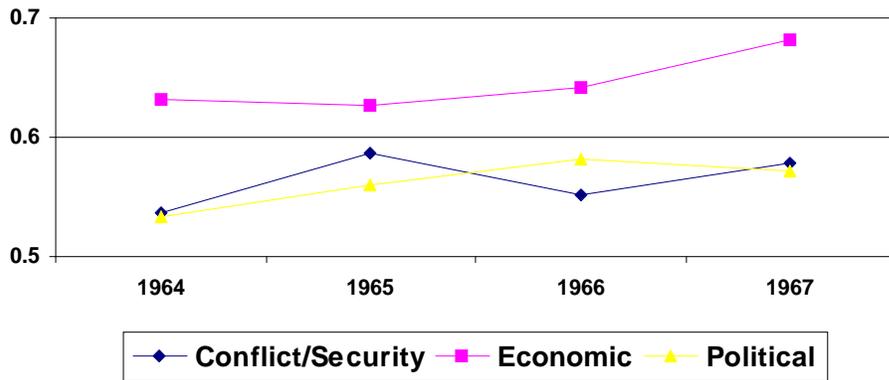
TASK for Clinton by YEAR



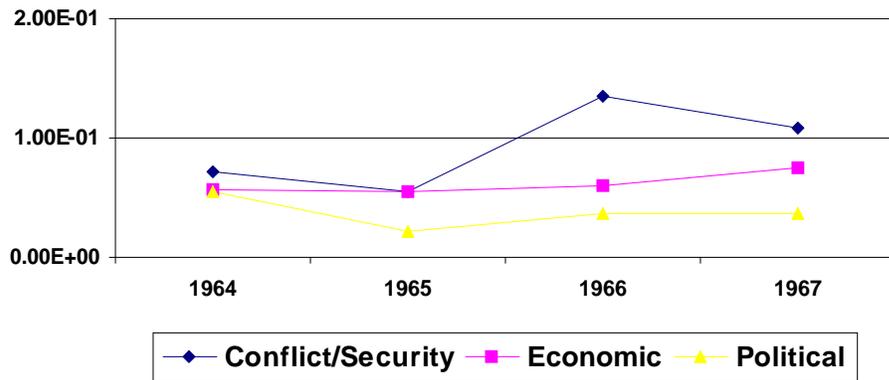
BACE for Johnson by YEAR



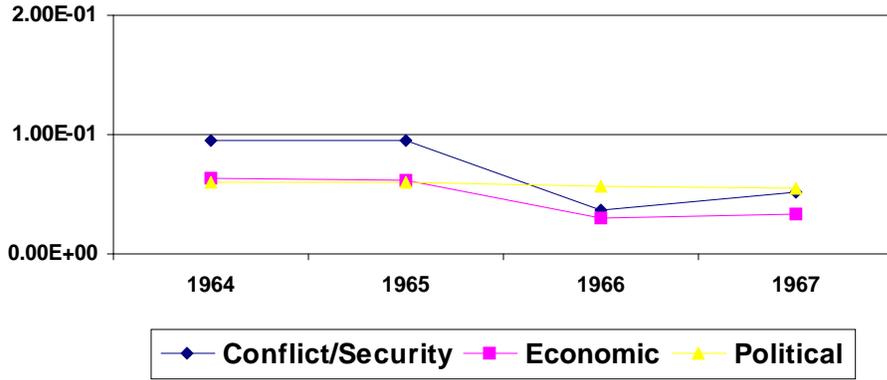
CC for Johnson by YEAR



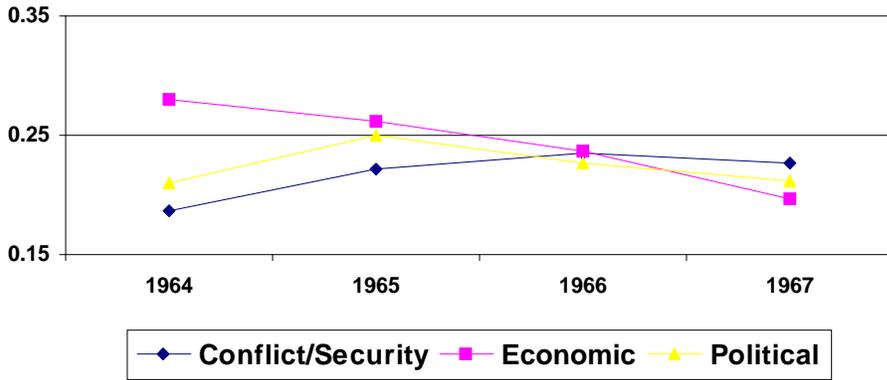
DIS for Johnson by YEAR



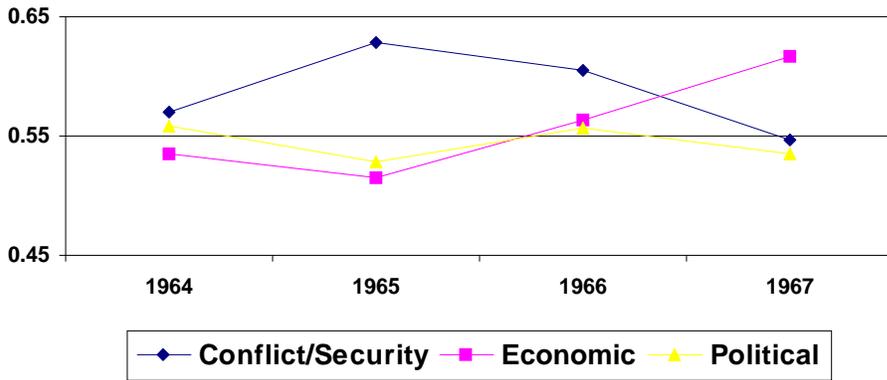
IGB for Johnson by YEAR



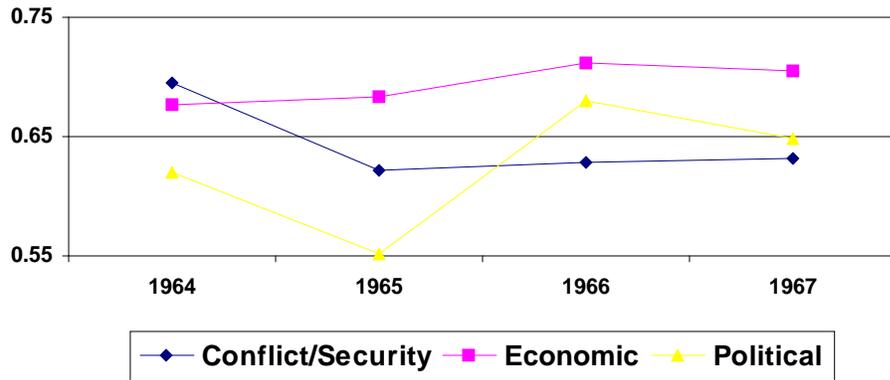
PWR for Johnson by YEAR



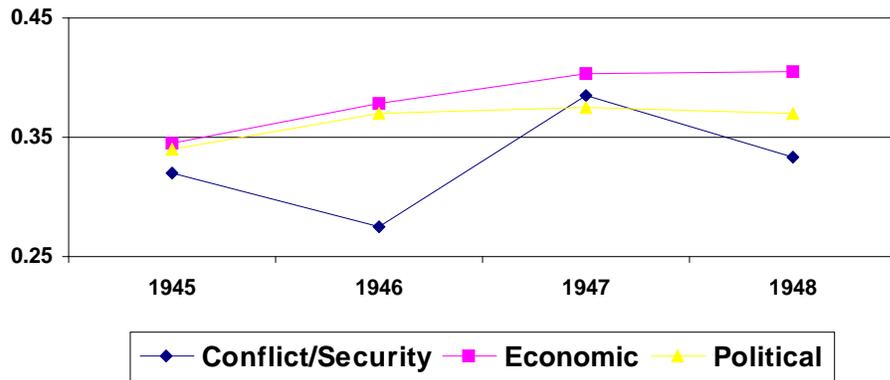
SC for Johnson by YEAR



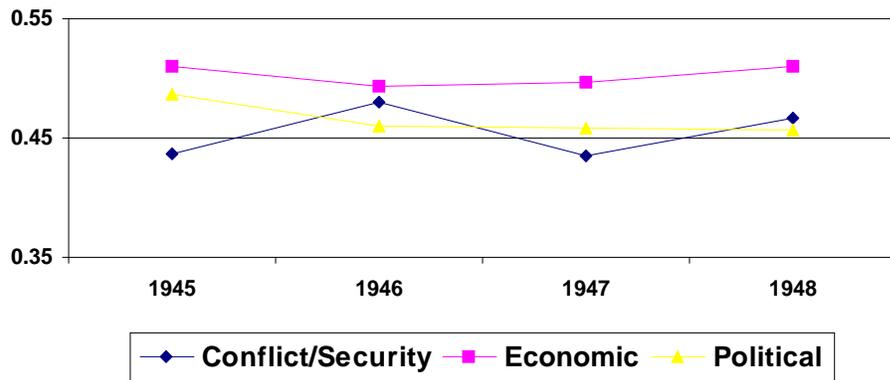
TASK for Johnson by YEAR



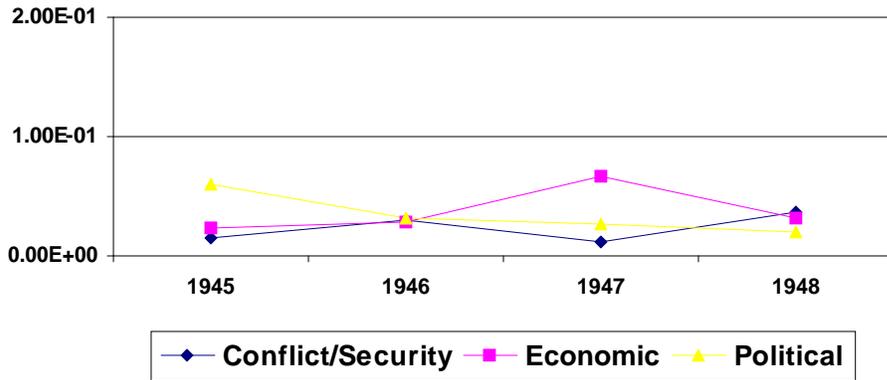
BACE for Truman by YEAR



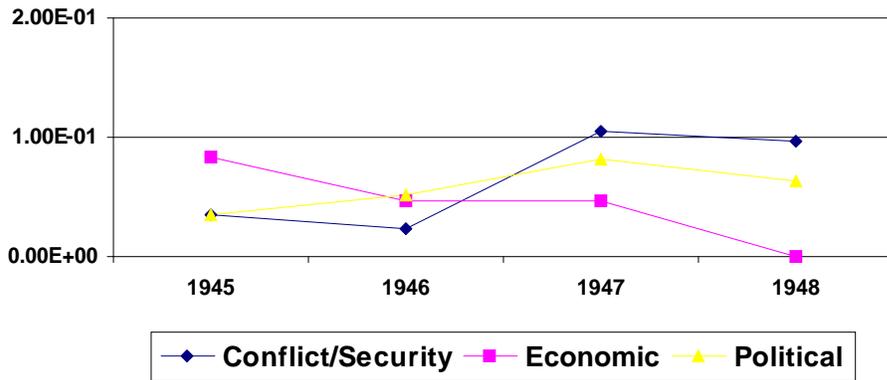
CC for Truman by YEAR



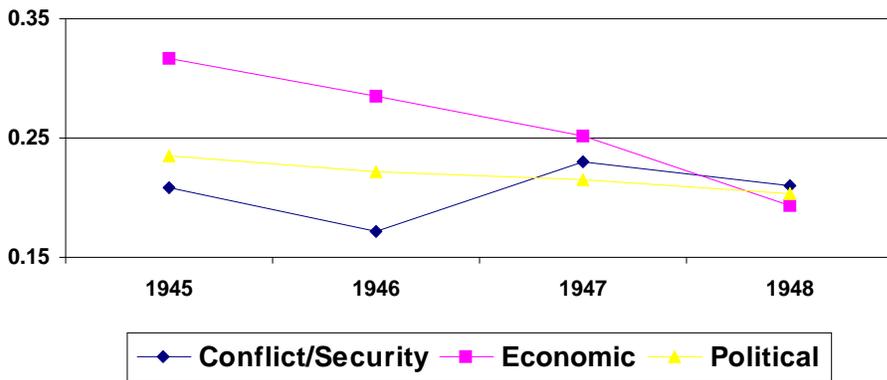
DIS for Truman by YEAR



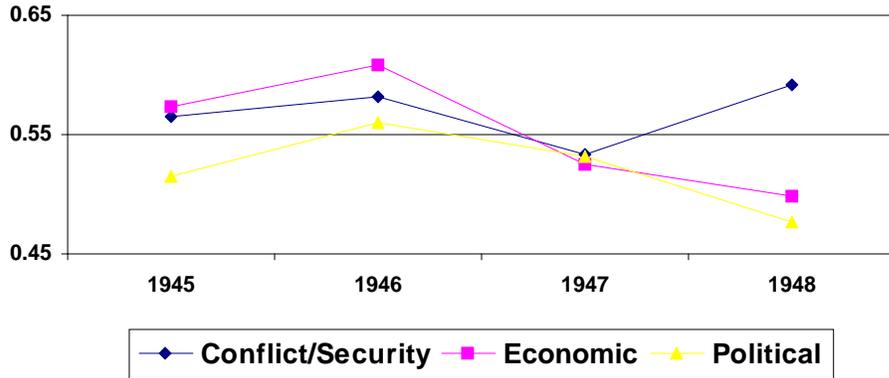
IGB for Truman by YEAR



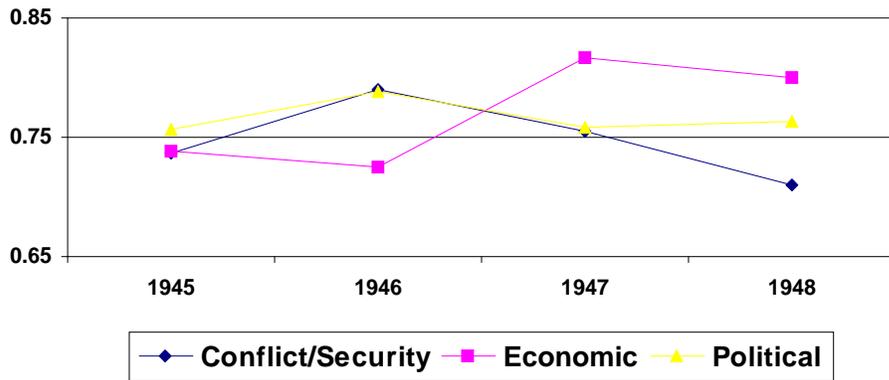
PWR for Truman by YEAR



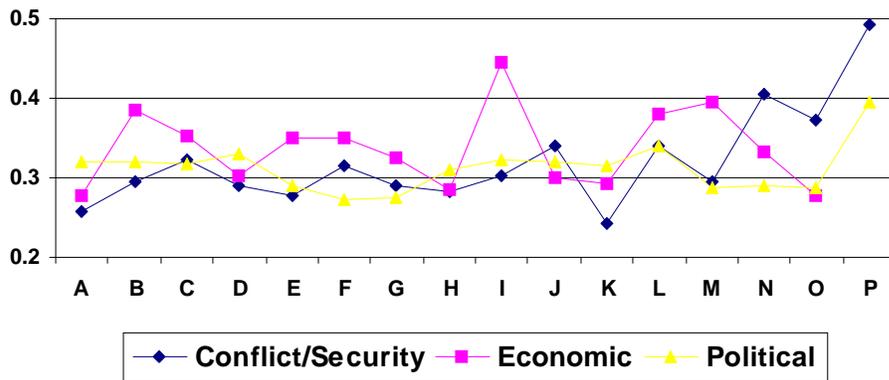
SC for Truman by YEAR



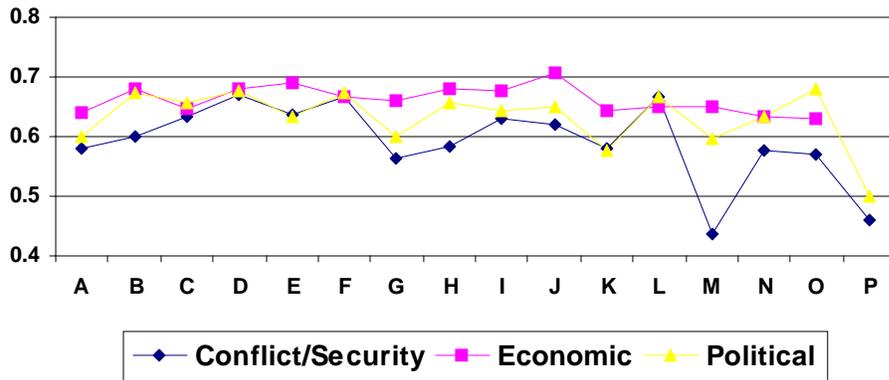
TASK for Truman by YEAR



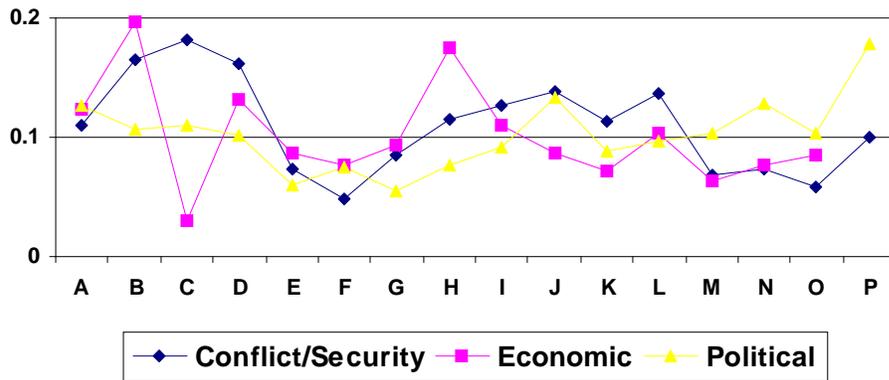
BACE for Bush by QTR



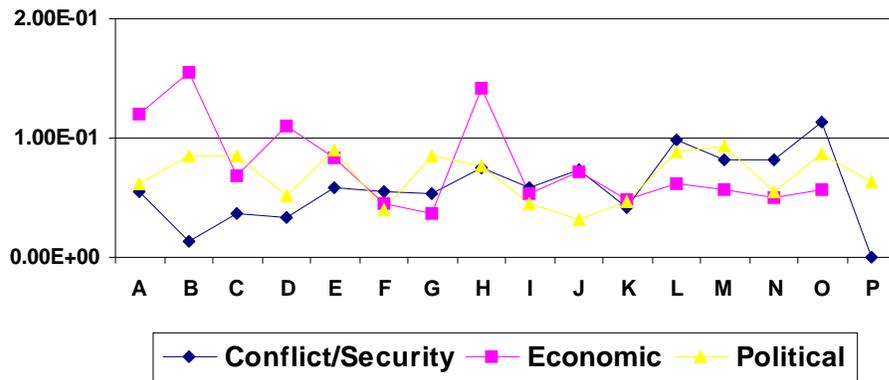
CC for Bush by QTR



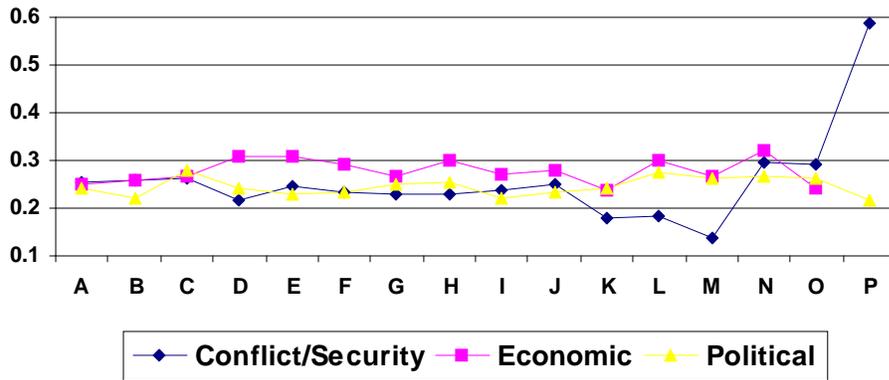
DIS for Bush by QTR



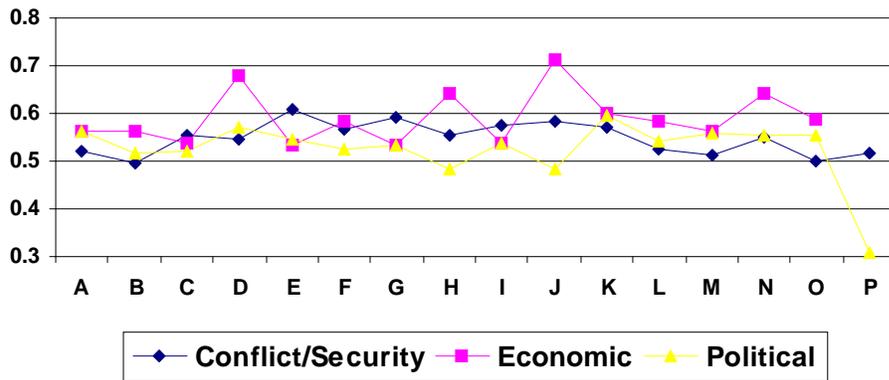
IGB for Bush by QTR



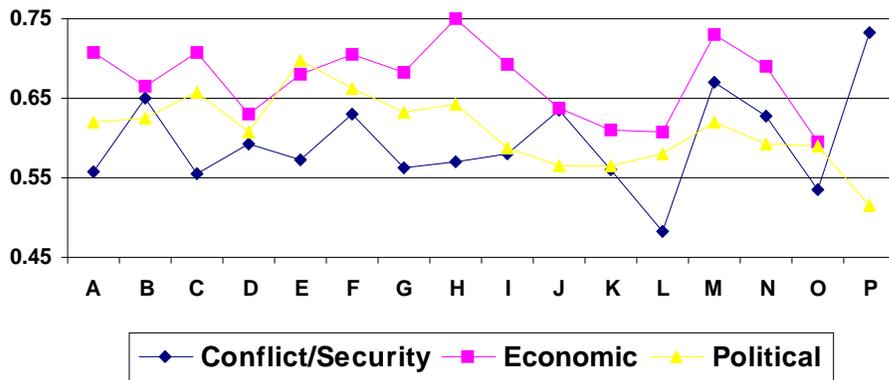
PWR for Bush by QTR



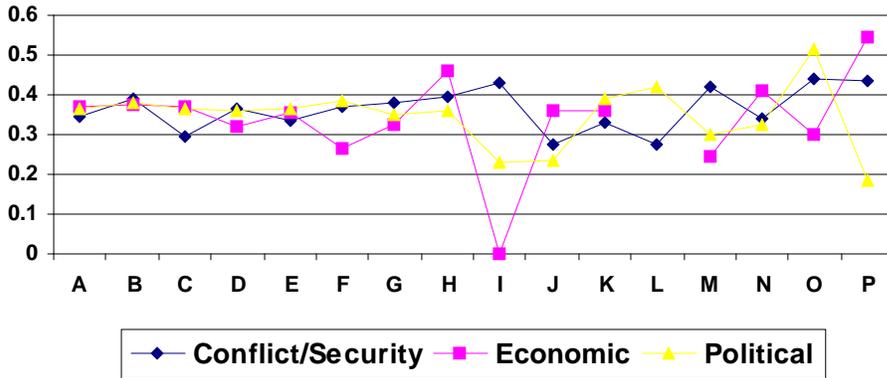
SC for Bush by QTR



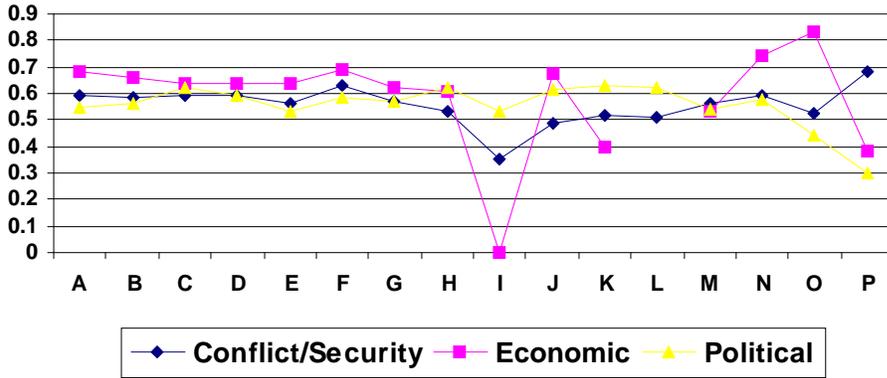
TASK for Bush by QTR



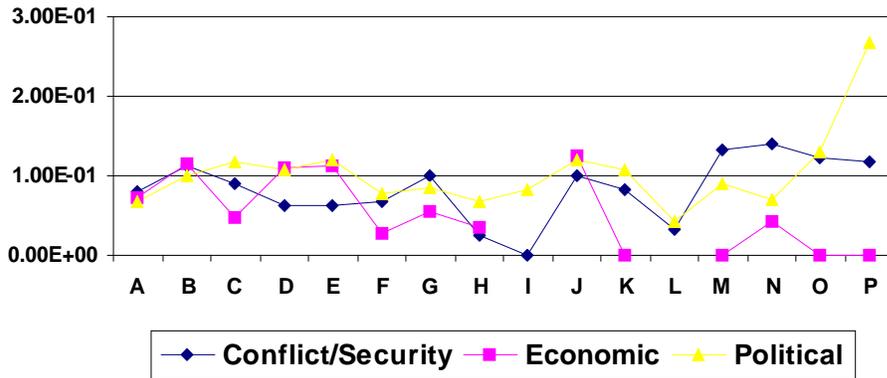
BACE for Clinton by QTR



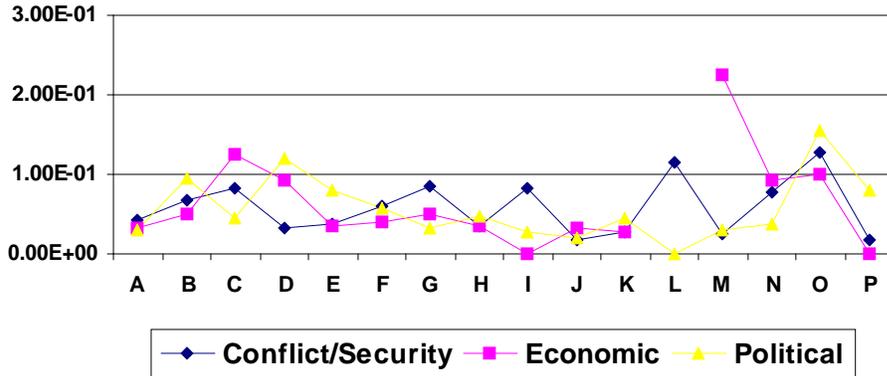
CC for Clinton by QTR



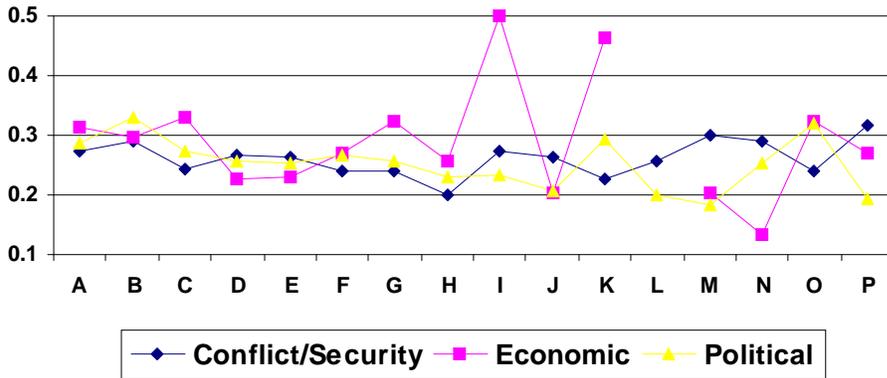
DIS for Clinton by QTR



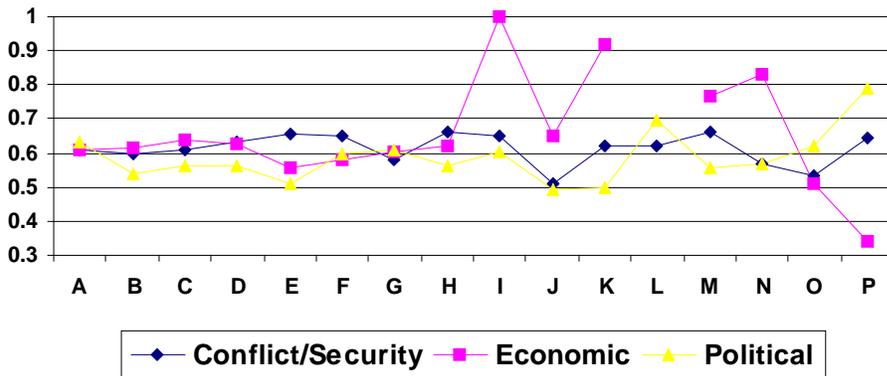
IGB for Clinton by QTR



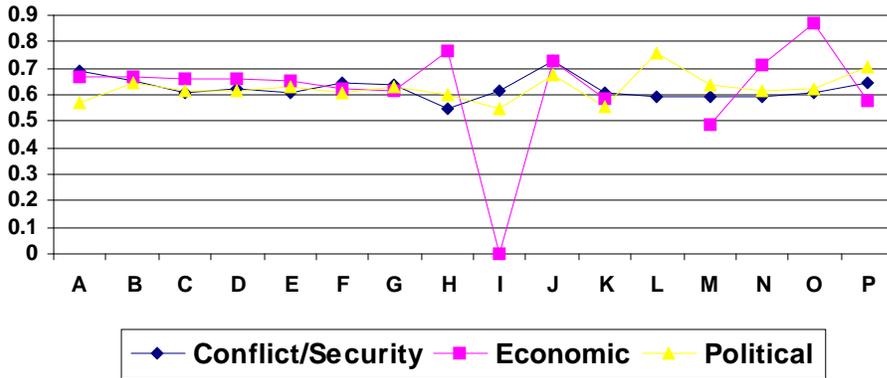
PWR for Clinton by QTR



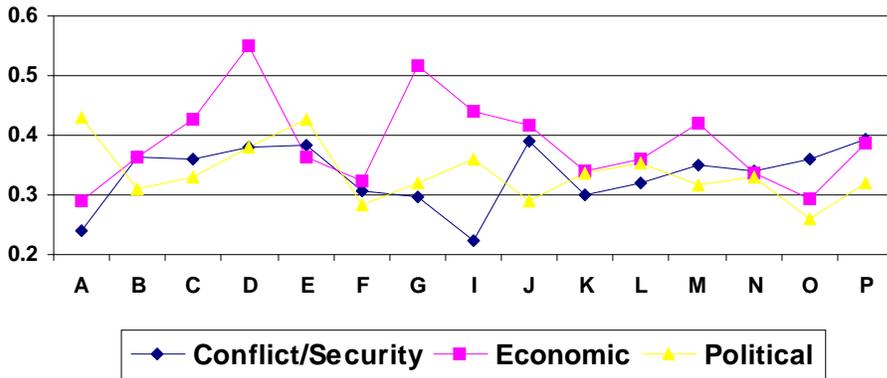
SC for Clinton by QTR



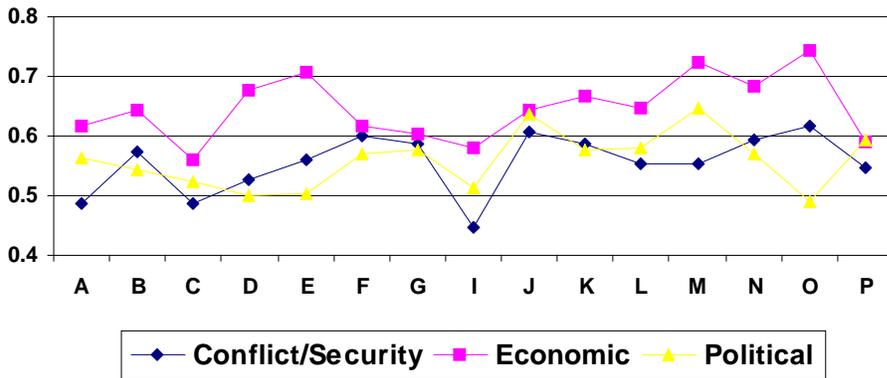
TASK for Clinton by QTR



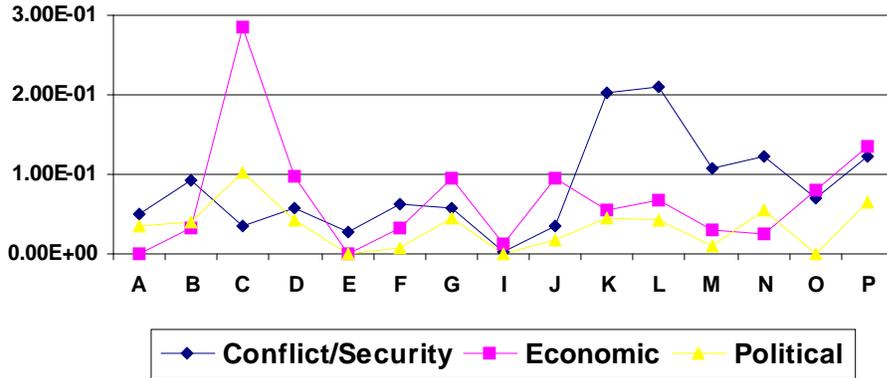
BACE for Johnson by QTR



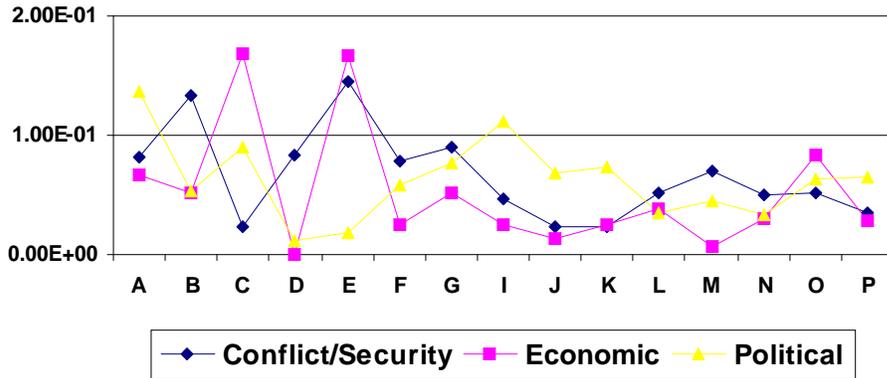
CC for Johnson by QTR



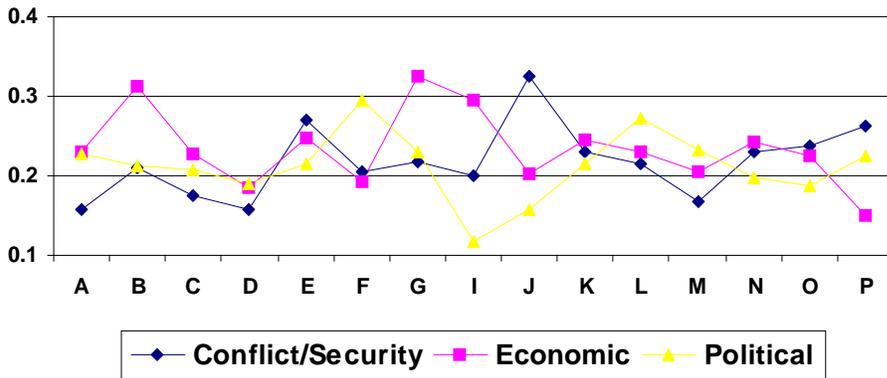
DIS for Johnson by QTR



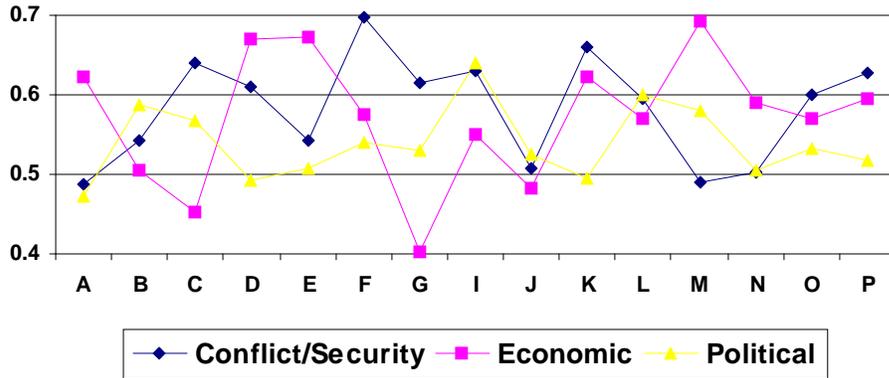
IGB for Johnson by QTR



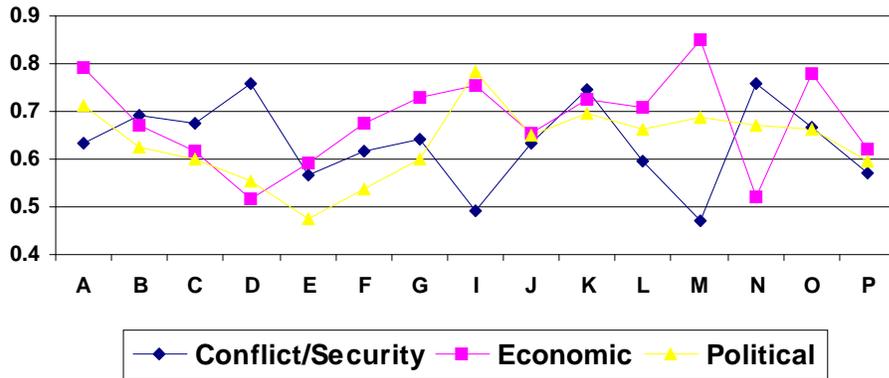
PWR for Johnson by QTR



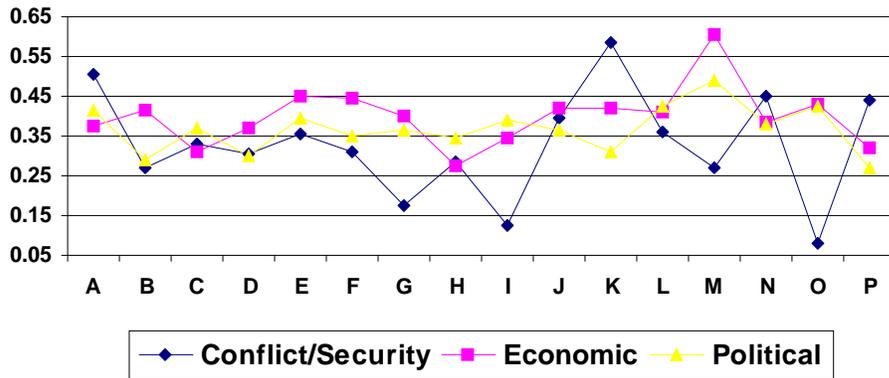
SC for Johnson by QTR



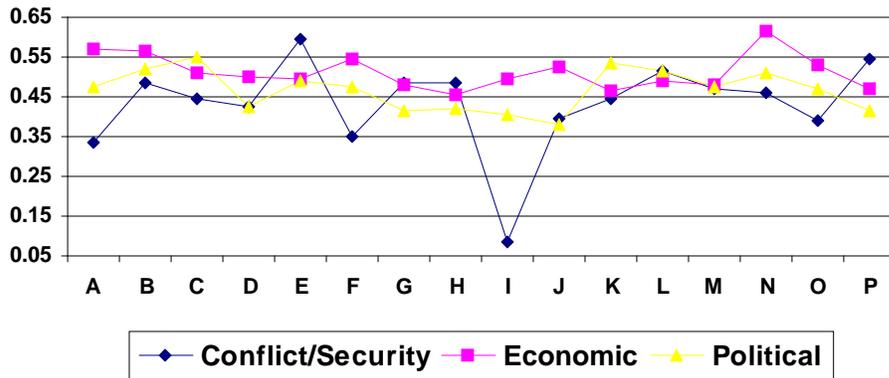
TASK for Johnson by QTR



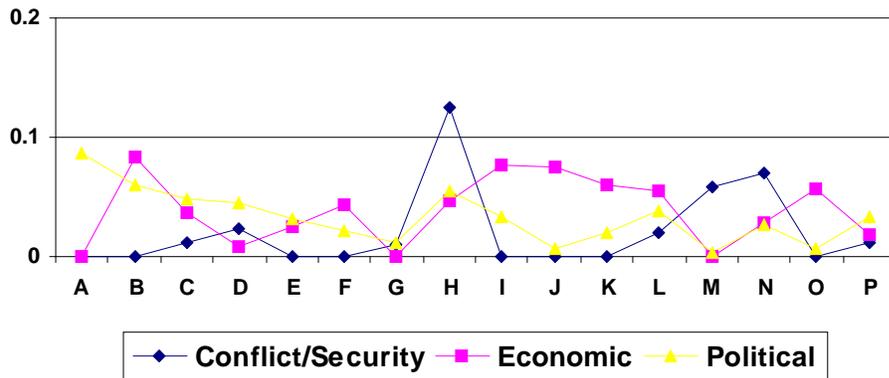
BACE for Truman by QTR



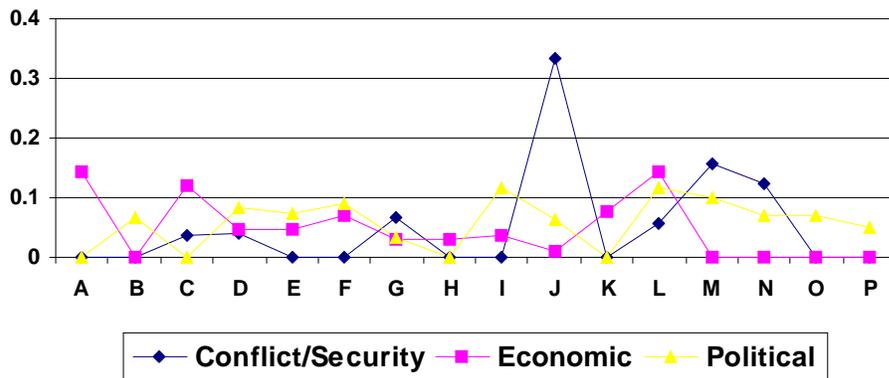
CC for Truman by QTR



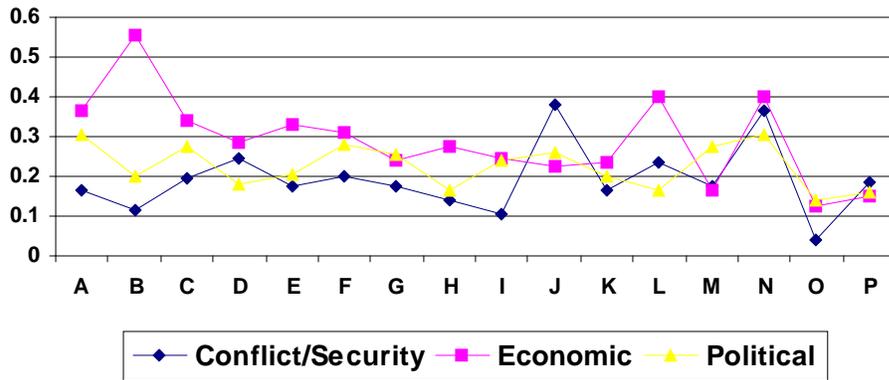
DIS for Truman by QTR



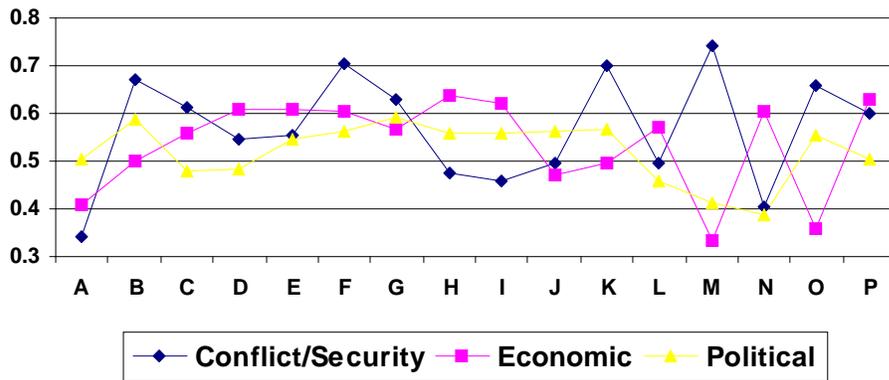
IGB for Truman by QTR



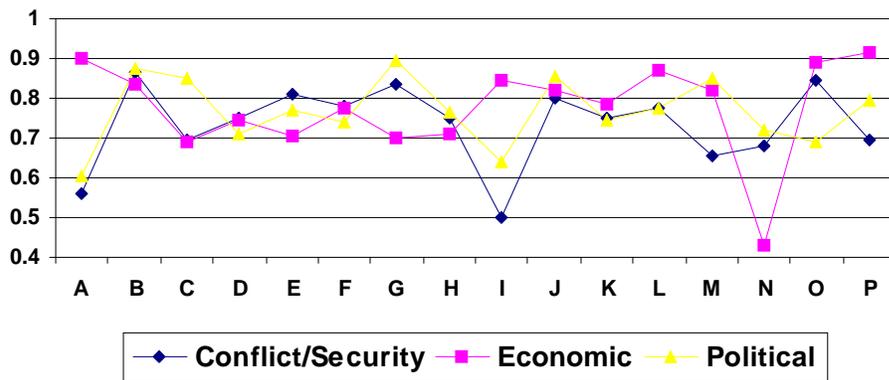
PWR for Truman by QTR



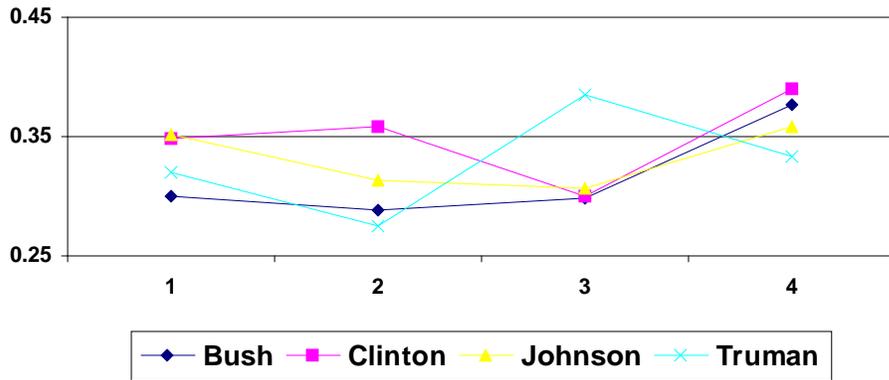
SC for Truman by QTR



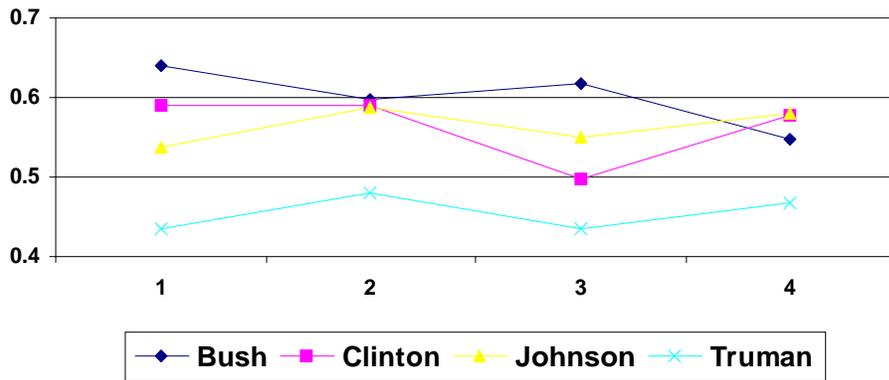
TASK for Truman by QTR



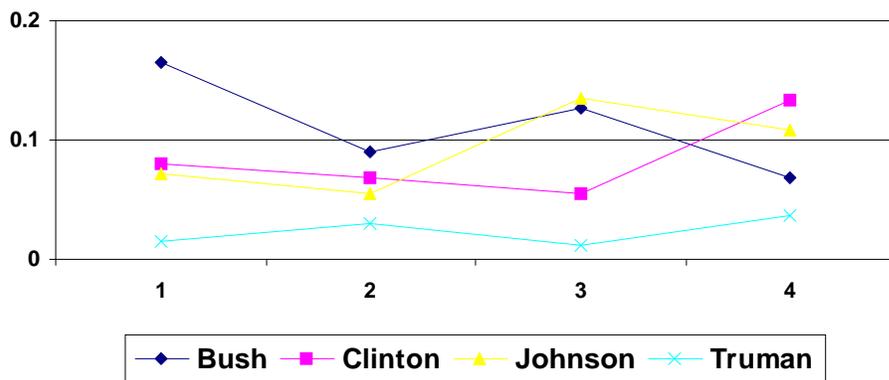
BACE for Conflict/Security by YEAR



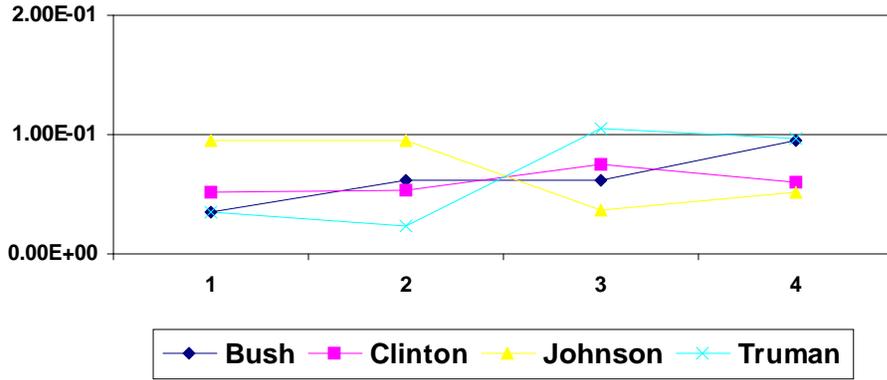
CC for Conflict/Security by YEAR



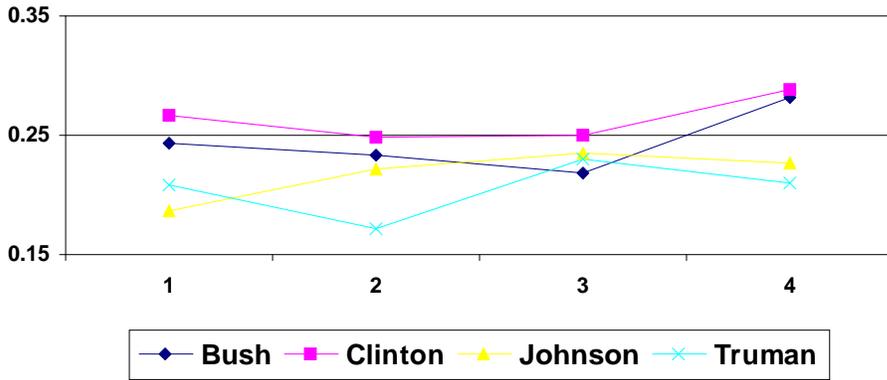
DIS for Conflict/Security by YEAR



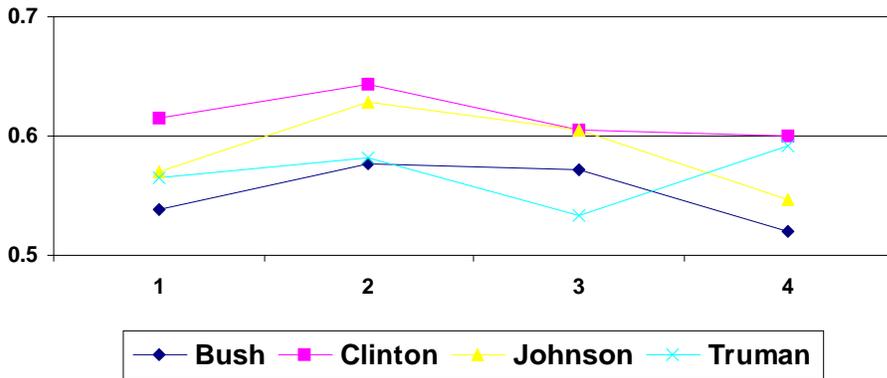
IGB for Conflict/Security by YEAR



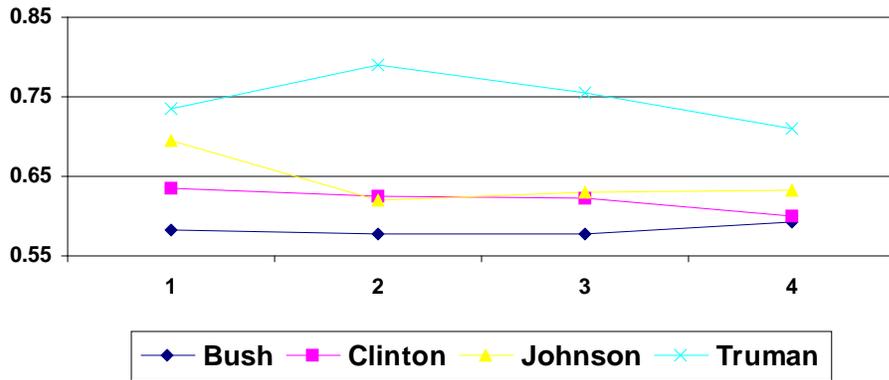
PWR for Conflict/Security by YEAR



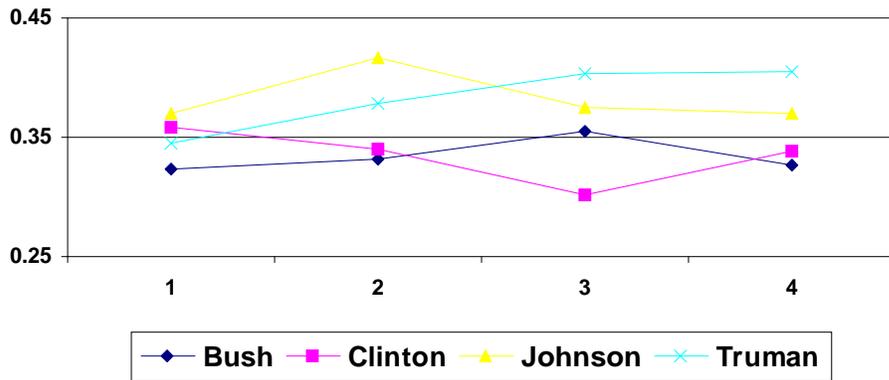
SC for Conflict/Security by YEAR



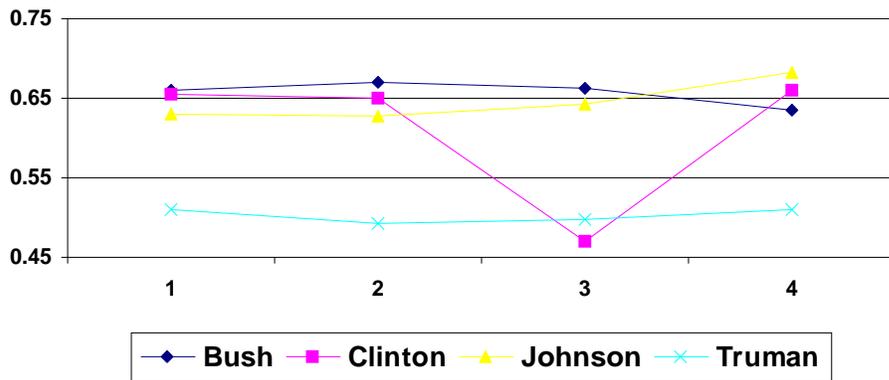
TASK for Conflict/Security by YEAR



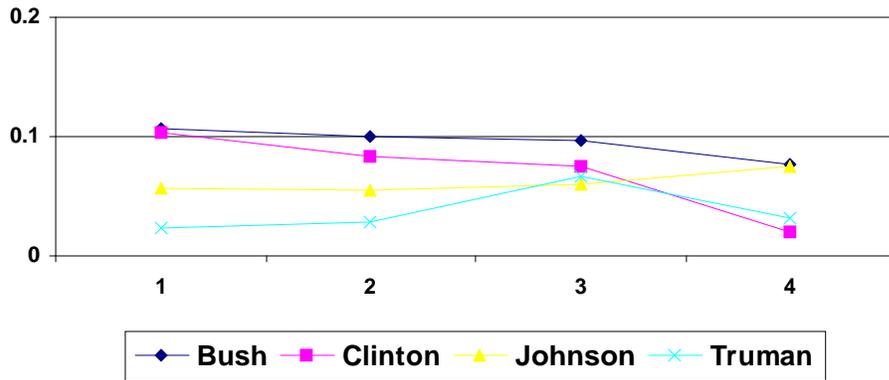
BACE for Economics by YEAR



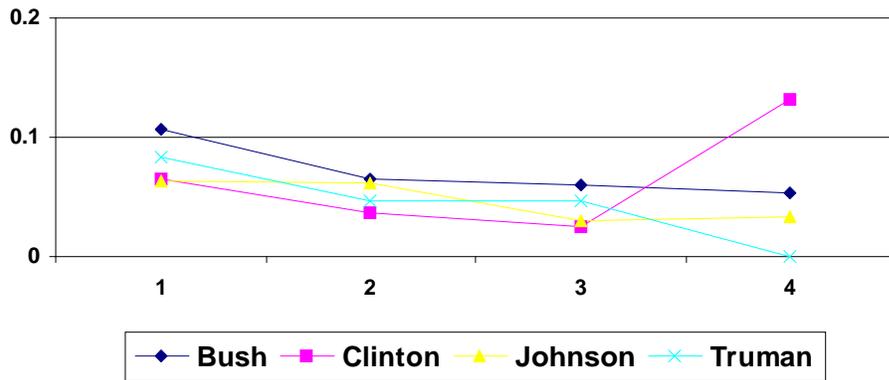
CC for Economics by YEAR



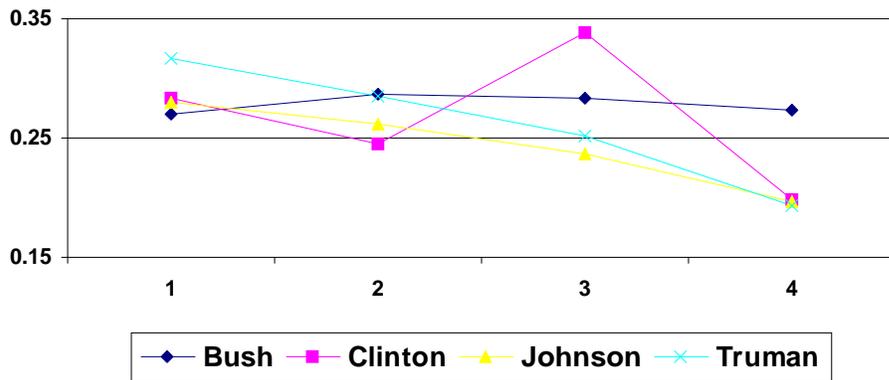
DIS for Economics by YEAR



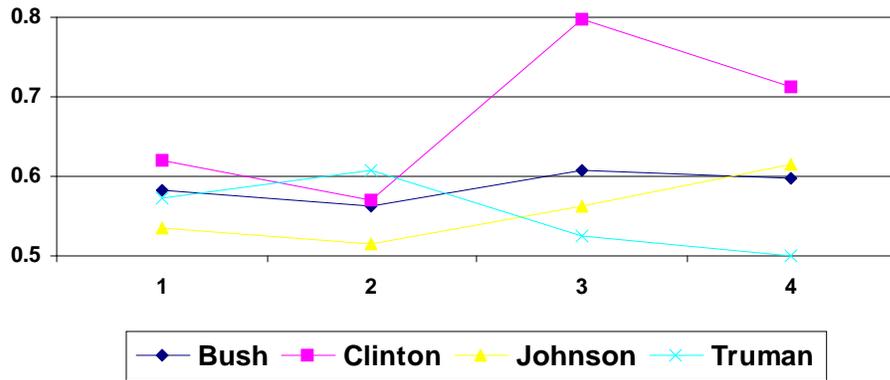
IGB for Economics by YEAR



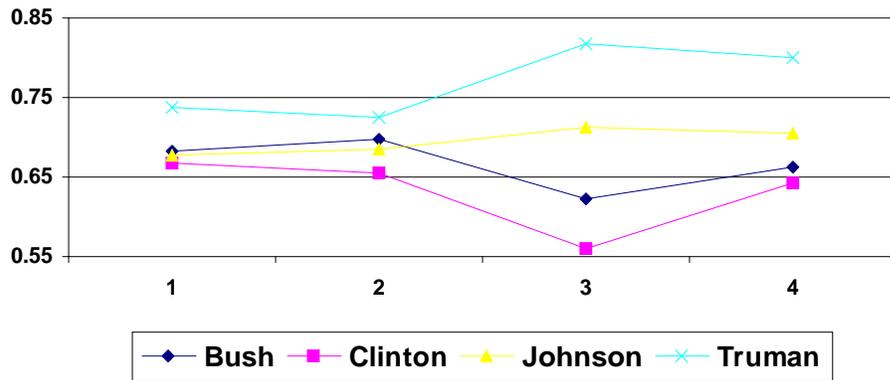
PWR for Economics by YEAR



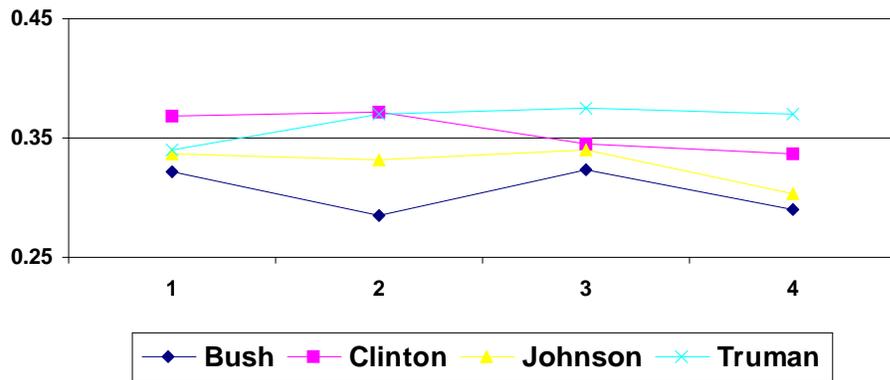
SC for Economics by YEAR



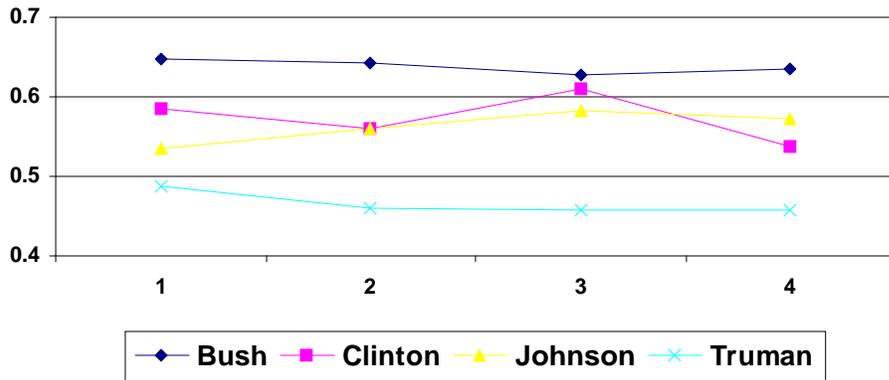
TASK for Economics by YEAR



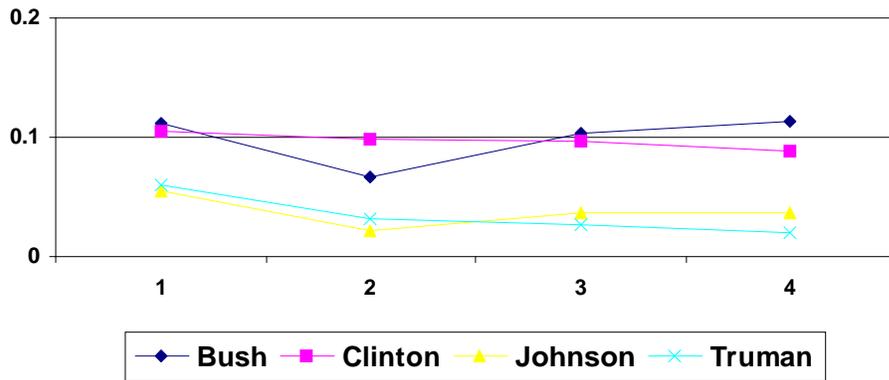
BACE for Political by YEAR



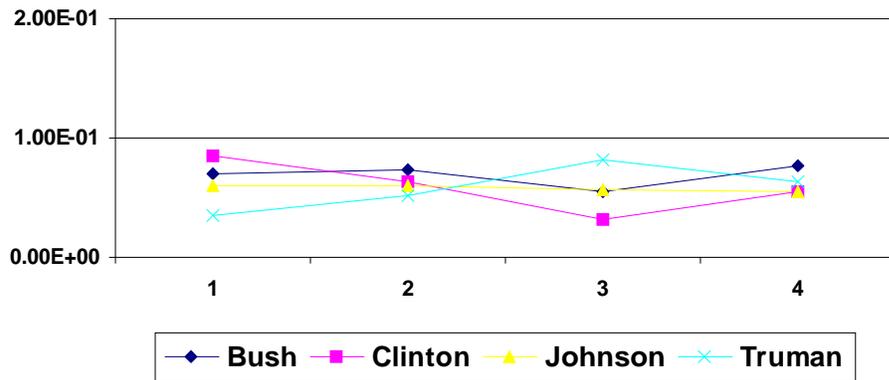
CC for Political by YEAR



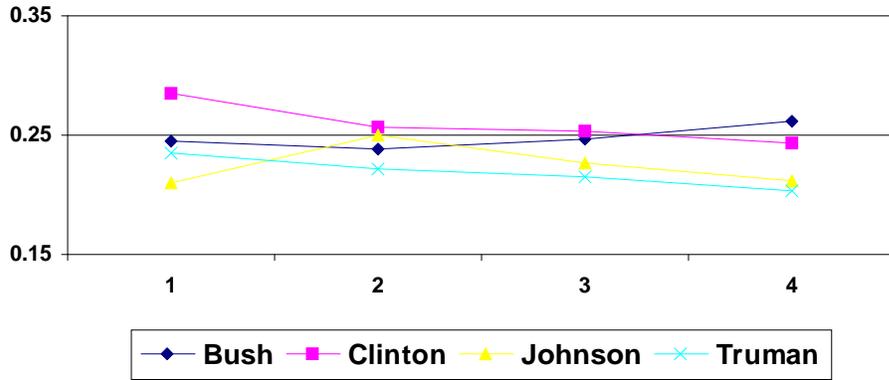
DIS for Political by YEAR



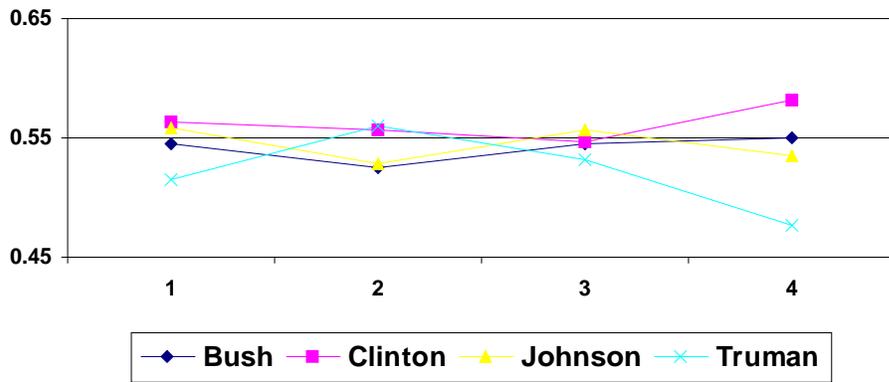
IGB for Political by YEAR



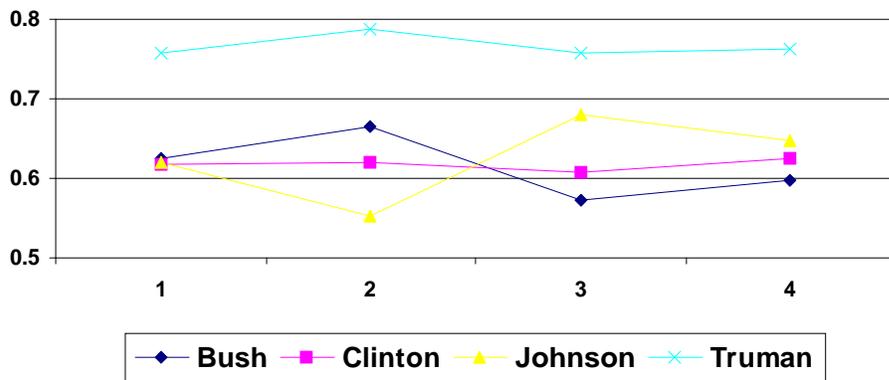
PWR for Political by YEAR



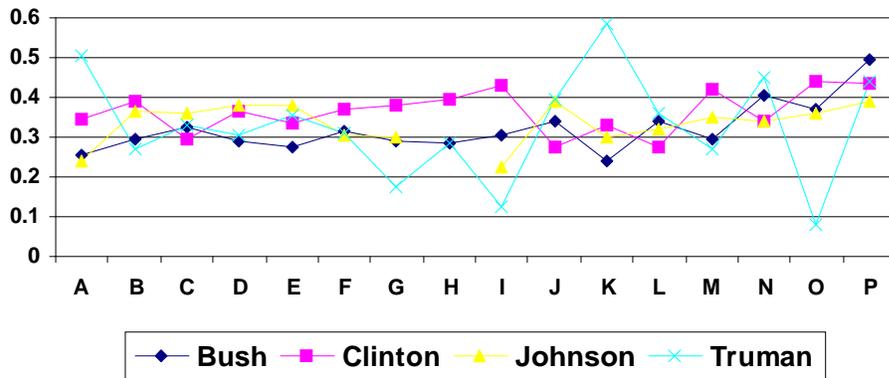
SC for Political by YEAR



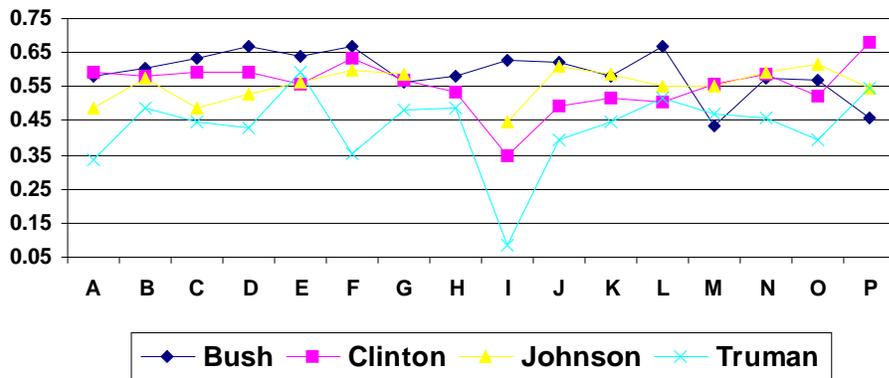
TASK for Political by YEAR



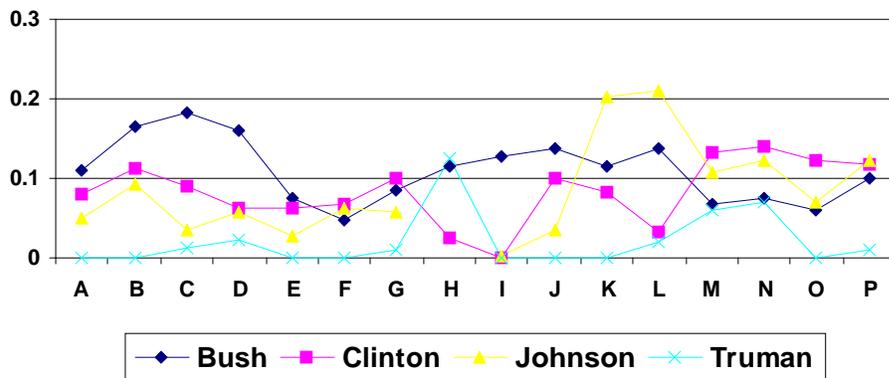
BACE for Conflict/Security by QTR



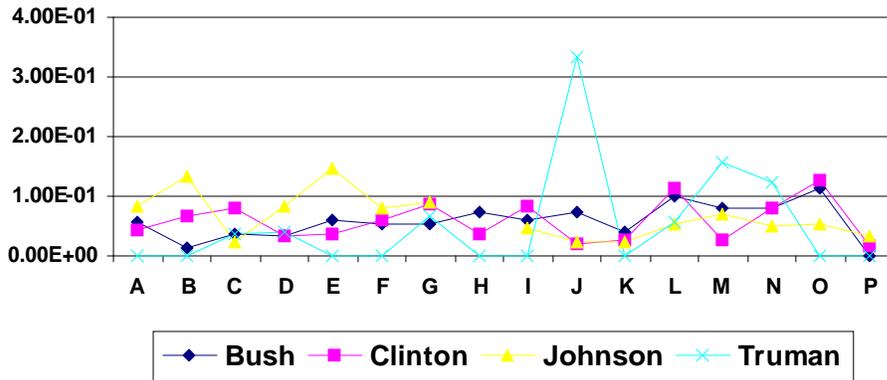
CC for Conflict/Security by QTR



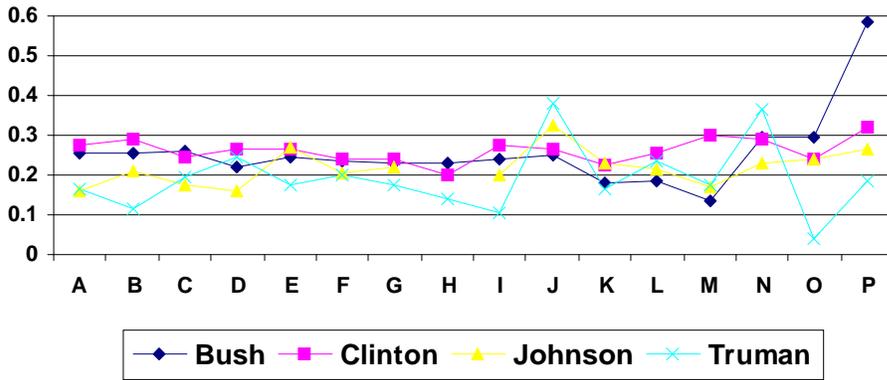
DIS for Conflict/Security by QTR



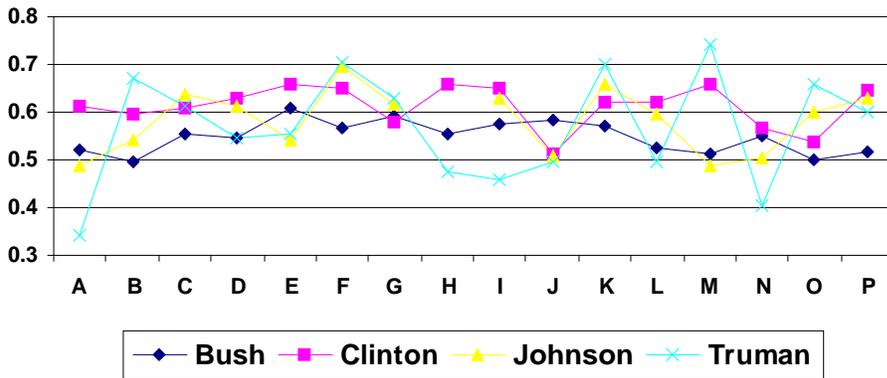
IGB for Conflict/Security by QTR



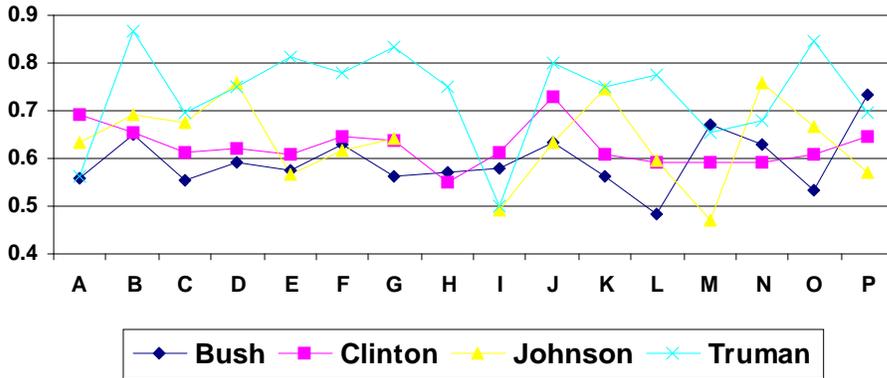
PWR for Conflict/Security by QTR



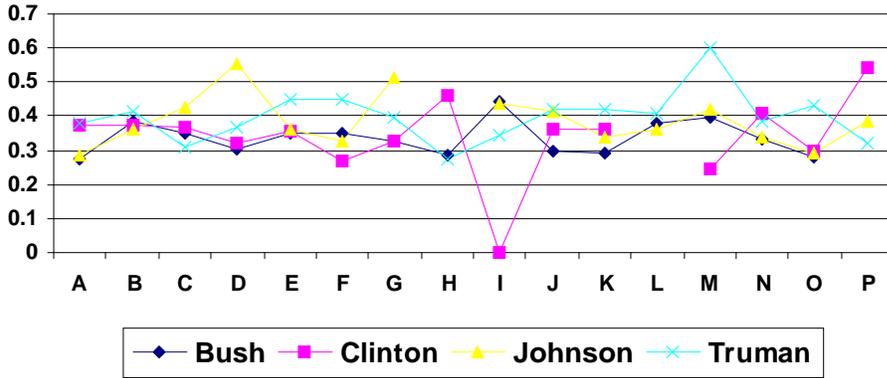
SC for Conflict/Security by QTR



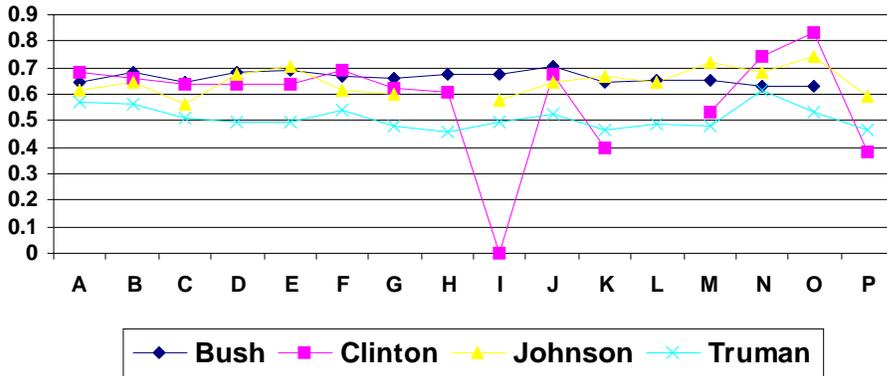
TASK for Conflict/Security by QTR



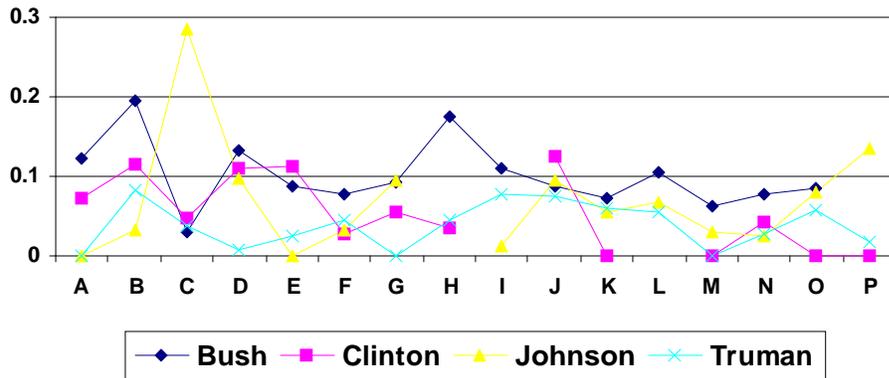
BACE for Economics by QTR



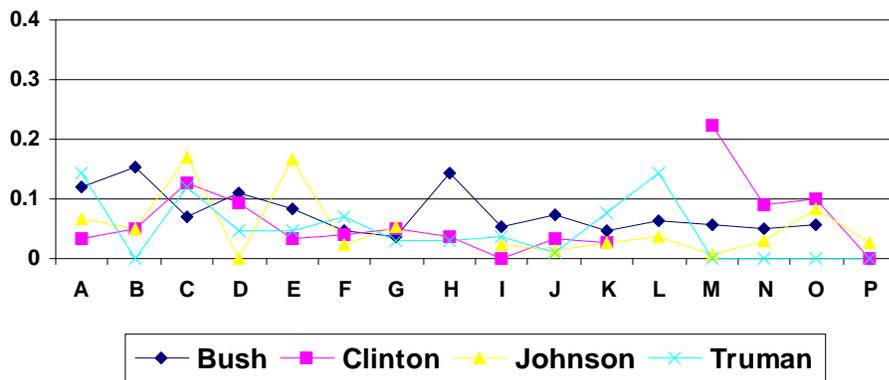
CC for Economics by QTR



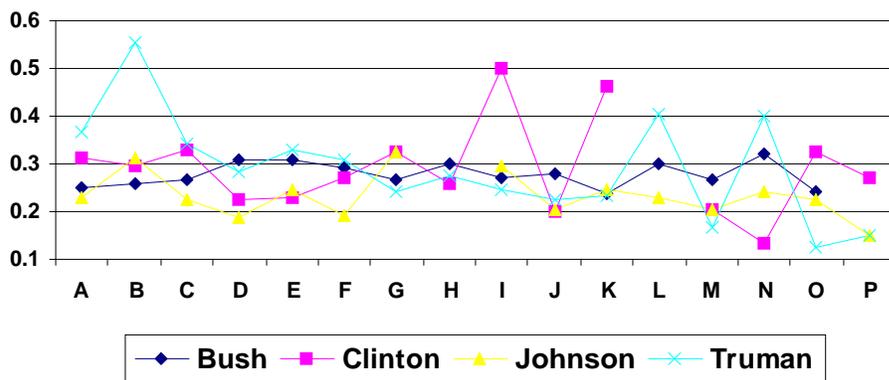
DIS for Economics by QTR



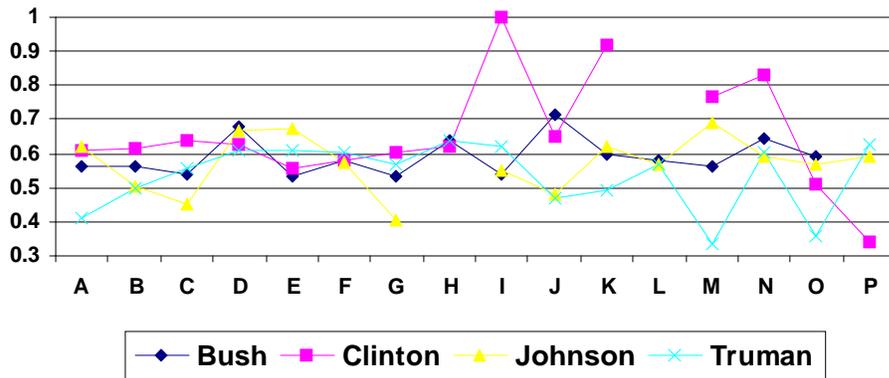
IGB for Economics by QTR



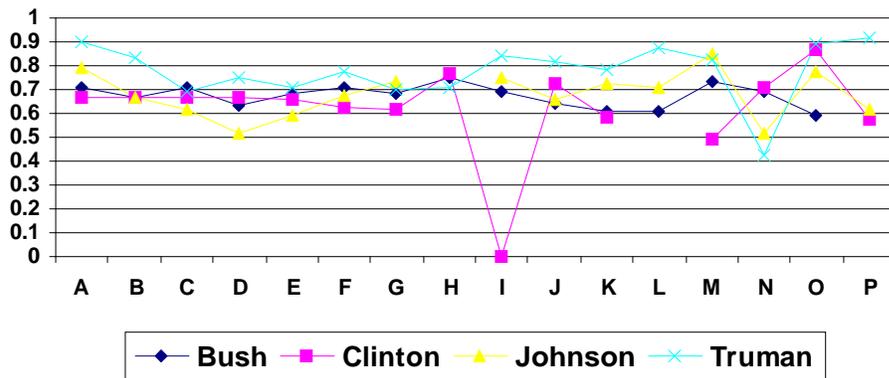
PWR for Economics by QTR



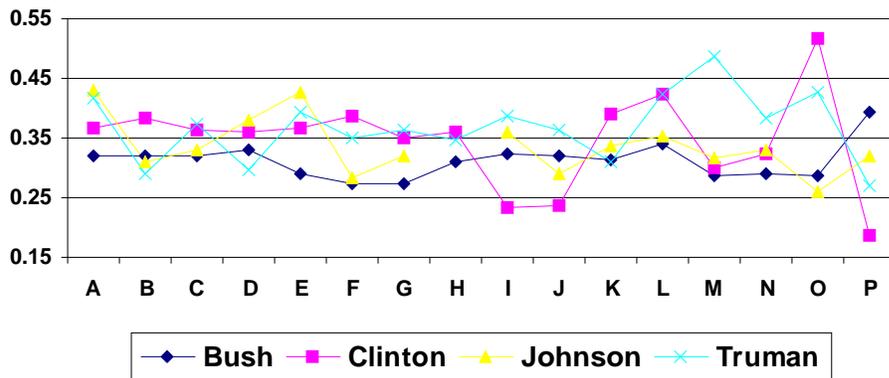
SC for Economics by QTR



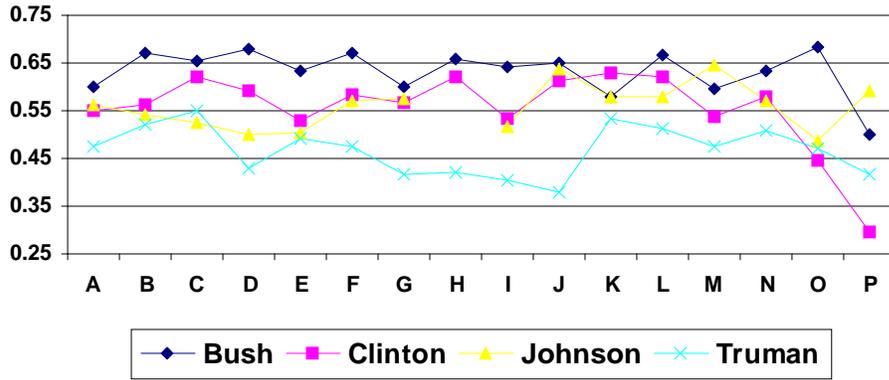
TASK for Economics by QTR



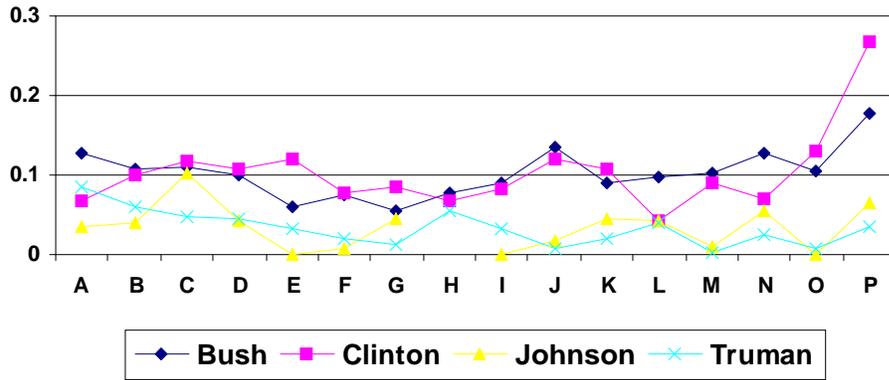
BACE for Political by QTR



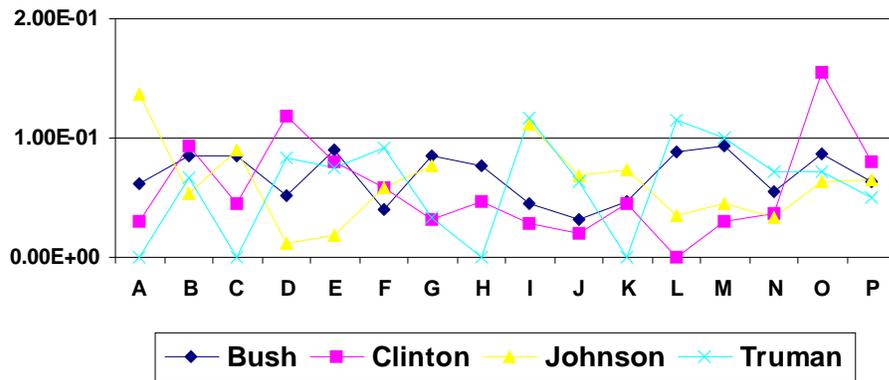
CC for Political by QTR



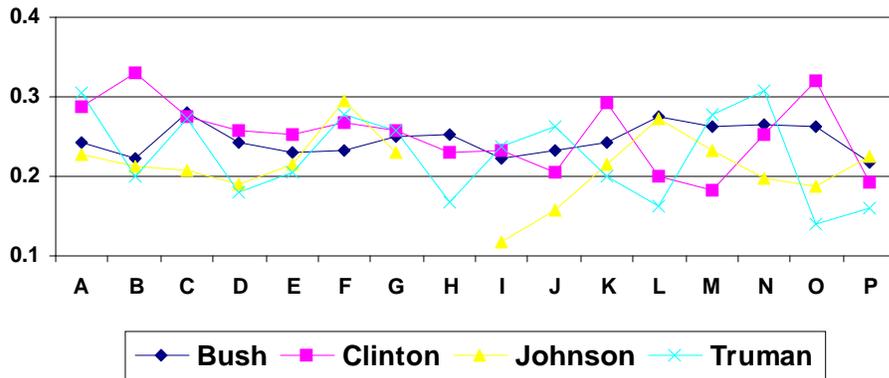
DIS for Political by QTR



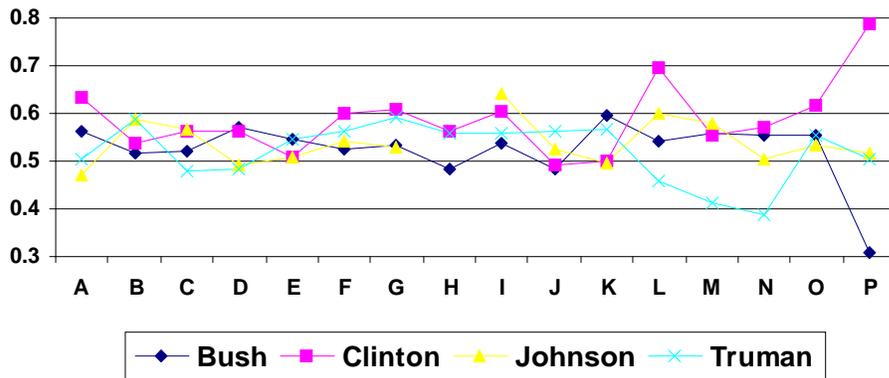
IGB for Political by QTR



PWR for Political by QTR



SC for Political by QTR



TASK for Political by QTR

