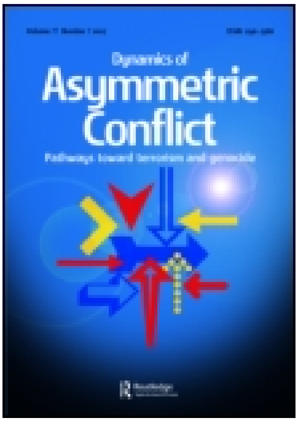

Dynamics of Asymmetric Conflict
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rdac20

The terrorist rhetorical style and its consequences for understanding terrorist violence

Lucian Gideon Conway III [a] & Kathrene R. Conway [a]
[a] The University of Montana, Missoula, MT, USA

Available online: 17 Nov 2011

Dynamics of Asymmetric Conflict
Vol. 4, No. 2, July 2011, 175–192

Routledge
Taylor & Francis Group

The terrorist rhetorical style and its consequences for understanding terrorist violence

Lucian Gideon Conway III* and Kathrene R. Conway

The University of Montana, Missoula, MT, USA

(*Received 10 March 2011; final version received 27 September 2011*)

The linguistic content analyses systems used in the present special issue covered a large range of conceptual and methodological approaches from many of the major researchers in the field. Taking a broad view of this massive set of results, what can we learn about terrorist groups' rhetoric? We present here a summary set of analyses that suggests terrorist rhetoric (compared to non-terrorist control groups) was more social in nature, contained more rhetoric reflecting control and power, and was consistently lower on multiple measurements of rhetoric complexity. Further, analyses on automated systems suggested that the typical terrorist style of rhetoric became even more exaggerated among terrorist groups as an impending violent attack by their group neared. This study highlights the benefits of conducting linguistic content analyses of terrorist groups' public rhetoric. Applied rigorously, this method can contribute to our understanding of how these groups differ from their non-terrorist counterparts and potentially provide indicators of when they may be ready to engage in further violent activity.

Keywords: terrorism; rhetoric; content analysis

The papers presented in this special issue have provided evidence that we can use linguistic content analytic (LCA) techniques both to (1) distinguish terrorist groups from ideologically similar non-terrorist groups, and (2) predict when terrorist groups will engage in violence. Here, in our concluding article, we hope to tie this diverse set of papers together.

In particular, we aim to provide a broad theoretical summary of the various systems presented here, focusing on similarities and differences in results across systems. To do so, we first summarize prior results comparing terrorist and non-terrorist group rhetoric at a ''dimension'' level rather than at a ''system'' level. In other words, we present results here focusing on constructs that overlap across systems (e.g. multiple systems have measurements relevant to complexity).

This article is a bit more than just a descriptive re-capitulation, however. Indeed, we here present some new analyses (as well as some analyses that overlap quite a bit with those in the prior papers).[1] Our larger goal here is to attempt to define a broad terrorist style of rhetoric that cuts across specific LCA systems, and then see if this terrorist rhetorical style can help us predict when terrorist groups will engage in violence.

*Corresponding author. Email: Luke.conway@umontana.edu

ISSN 1746-7586 print/ISSN 1746-7594 online

http://dx.doi.org/10.1080/17467586.2011.627940
http://www.tandfonline.com

In doing so, we tackle two additional issues that come with the diverse approaches used by the various researchers on the project. First, although working from the same data set, each research group used methods that were specific to their own LCA system. While this approach allowed each group to tailor their analyses to the strengths of their own system, it also makes direct cross-system comparison more difficult. For example, different researchers occasionally used different exclusion criteria for which documents would be included/excluded in their sample. While this individualized approach has its strengths, it also causes potential stumbling blocks to understanding because it is unclear (for example) whether potential differences that emerge between systems occurred due to substantive meaningful divergences or to the fact that the researchers were using a partially non-overlapping set of documents. As a result, in the present article we use a single set of exclusion/inclusion criteria that is applied equally across all the systems presented here. While this approach sometimes forces discrepant systems somewhat into the same mold, it also allows for cleaner comparisons across systems.

Second, there are multiple ways to approach the issue of *when* terrorists will attack. This issue was addressed in a variety of ways in the papers presented here (indeed, some of the research papers here did not directly address this question at all). As a result, here in this summary article, we present a more comprehensive account of our ability to predict terrorist violence, with an eye towards the larger picture of how these LCA systems fit together theoretically.

Differences between terrorist and non-terrorist group rhetoric: a dimension-level summary

Before tackling the larger and more difficult question of predicting when terrorist groups will engage in violence, we first aimed to define more broadly the qualities of terrorist rhetoric. All of the papers presented here dealt with this issue, and thus we do not spend time here on any system specifically. Rather, our goal here is to provide a statistical summary that (1) conceptually organizes across dimensions rather than systems, and (2) uses identical exclusion criteria. In doing so, the vast majority of our results below parallel or directly match those presented in the prior papers; thus, what we present here is not so much "new" as it is re-organized for understanding the larger conceptual issue of characterizing terrorist rhetoric writ large.

All primary results reported below were computed in a 2 (Terrorist Group: Yes versus No) $\times$ 2 (Context: Transnational versus Arabian Peninsula) ANOVA. Our focus is on two relevant details: (1) The main effect of terrorism, which helps answer our first focal question about differences between the rhetoric of terrorist groups and that of their non-terrorist counterparts. (2) The interaction between terrorism and context, which lets us know if the effect of terrorism differed in the two locations. To the degree that we find a main effect of terrorism and no interaction, this suggests that the rhetoric of both terrorist groups differed from that of their non-terrorist comparison groups in roughly the same way. An interaction suggests that the effect of terrorism operated differently in the two contexts. (We do not focus here on main effect differences across contexts. Although interesting, such differences are beyond our present scope of understanding the effect of terrorism specifically.)

Below, we present the highlights of these analyses from our 2 $\times$ 2 framework, summarized at the level of each conceptual dimension. The interested reader can find more detailed summaries of each dimension in Tables 1–6. All differences discussed were significant at least at the $p < .05$ level, unless otherwise noted.[2]

Table 1. Domain-general complexity: terrorist versus non-terrorist rhetoric.

System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
DICTION	Complexity	4.60	4.51	4.95	4.54	*4.78*	*4.53*	***	***
Int. Complexity	Integrative Complexity	2.21	1.90	2.06	1.67	*2.14*	*1.79*	***	
Int. Complexity	Dialectical Complexity	1.29	1.27	1.30	1.19	*1.29*	*1.23*	ˆ	ˆ
Int. Complexity	Elaborative Complexity	2.03	1.71	1.85	1.53	*1.94*	*1.62*	***	
Int. Complexity	Defensive Complexity	0.74	0.44	0.53	0.34	*0.65*	*0.39*	***	
LIWC –summary variable	Complexity	0.58	–0.79	1.51	–1.22	*1.05*	*−1.02*	***	ˆ
LTA	Complexity	0.61	0.58	0.62	0.63	*0.62*	*0.60*	ˆ	

***$p < .001$; **$p < .01$; *$p < .05$; ˆ$p < .10$.

Table 2. Affiliation/interpersonal motives: terrorist versus non-terrorist rhetoric.

System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
DICTION	Rapport	1.74	3.08	2.52	3.30	*2.14*	*3.19*	***	
DICTION – summary	Commonality	49.89	50.29	50.52	50.28	*50.21*	*50.28*		
Motive imagery	Affiliation	–0.25	0.07	−0.40	0.40	*−0.33*	*0.23*	***	*
Motive imagery	Ingroup aff	0.61	0.97	0.12	1.37	*0.34*	*1.16*	***	**
Motive imagery	Outgroup aff	0.01	0.07	0.04	0.03	*0.03*	*0.06*		
SVS	Universalism	0.04	0.07	0.00	0.02	*0.03*	*0.04*		
LIWC – partial sum.	Social processes	8.57	12.15	8.00	12.71	*8.29*	*12.43*	***	ˆ

***$p < .001$; **$p < .01$; *$p < .05$; ˆ$p < .10$.

Table 3. Control/power: terrorist versus non-terrorist rhetoric.

System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
LTA	Control	0.29	0.35	0.27	0.38	*0.28*	*0.37*	**	
SVS	Self-direction	0.05	0.07	0.00	0.04	*0.04*	*0.06*		
VICS	Political future predictable	0.11	0.14	0.14	0.16	*0.13*	*0.15*	**	
VICS	Historical control	0.02	0.10	0.04	0.11	*0.03*	*0.10*	***	
VICS	Role of chance	1.00	0.99	0.99	0.98	*100.00*	*98.00*	***	
Motive Imagery	Power	–0.28	0.08	–0.24	0.23	*–0.26*	*0.16*	***	
LIWC – summary	Status	–0.35	0.13	–0.64	0.54	*–0.49*	*0.34*	***	ˆ
LTA	Power	0.21	0.23	0.23	0.23	*0.22*	*0.23*		
SVS	Power	0.09	0.09	0.50	0.05	*0.14*	*0.06*	***	***

*** $p < .001$; ** $p < .01$; * $p < .05$; ˆ $p < .10$.

Table 4. Aggression/distrust: terrorist versus non-terrorist rhetoric.

System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
DICTION	Aggression	10.23	10.34	9.74	9.21	*9.98*	*9.77*		
DICTION	Cooperation	3.42	4.08	4.26	3.07	*3.84*	*3.57*		**
SVS	Benevolence	0.03	0.07	0.00	0.08	*0.03*	*0.07*		
LTA	Ingroup bias	0.10	0.10	0.06	0.11	*0.08*	*0.11*	ˆ	ˆ
LTA	Distrust of others	0.24	0.27	0.27	0.27	*0.25*	*0.27*		
VICS	Strategy: cooperative	0.38	0.43	0.47	0.45	*0.42*	*0.44*		
VICS	Tactics: cooperative	0.16	0.13	0.16	0.12	*0.16*	*0.13*		
VICS	Political universe friendly	0.05	0.07	0.16	0.19	*0.11*	*0.13*		

*** $p < .001$; ** $p < .01$; * $p < .05$; ˆ $p < .10$.

Table 5. Achievement: terrorist versus non-terrorist rhetoric.

	Achievement								
System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
DICTION	Accomplish	10.14	7.58	13.78	9.70	*11.98*	*8.65*	***	
Motive Imagery	Achievement	–0.28	–0.07	–0.23	0.44	*–0.26*	*0.18*	***	*
LIWC	Work	2.02	1.51	2.48	1.41	*2.26*	*1.46*	***	**
LIWC	Achievement	2.29	2.02	2.43	2.28	*2.36*	*2.15*	*	
LIWC	Leisure	0.27	0.36	0.63	0.57	*0.45*	*0.46*		ˆ
LIWC	Money	0.31	0.50	0.72	0.32	*0.52*	*0.41*	ˆ	***
LTA	Task orient.	0.59	0.40	0.64	0.44	*0.61*	*0.42*	***	
SVS	Achievement	0.32	0.23	0.00	0.32	*0.27*	*0.28*		*

***$p < .001$; **$p < .01$; *$p < .05$; ˆ$p < .10$.

Table 6. Optimism/confidence: terrorist versus non-terrorist rhetoric.

	Optimism/confidence								
System	Variable	HT	AQ-C	MIRA	AQAP	*Non-terrorist*	*Terrorist*	Terror effect	Interaction
DICTION – summary	Optimism	49.38	49.41	48.34	49.70	*48.86*	*49.65*	*	
LTA	Self-confidence	0.28	0.16	0.32	0.17	*0.29*	*0.16*	**	
VICS	Political universe optimism	–0.01	–0.02	0.07	0.06	*0.03*	*0.02*		

***$p < .001$; **$p < .01$; *$p < .05$; ˆ$p < .10$.

Domain-general complexity

One clear pattern to emerge from this dimension-level view is that terrorist rhetoric is comprised of a simpler structure, at a general level, than non-terrorist rhetoric (this is consistent with prior research; see Smith, Suedfeld, Conway III, & Winter, 2008, Suedfeld & Leighton, 2002). This pattern was in evidence across a fairly large number of measures from both automated and manual systems, including significant effects from summative automated measures of complexity from LIWC, a single-item measurement of complexity from DICTION, and two of the three complexity measurements from the integrative complexity system. Further, marginally significant effects for LTA's complexity measurement and the third integrative complexity measurement (dialectical complexity) were also in the same direction. Overall, then, although there is variability among the strength of these effects, overwhelming similarity exists in the pattern: across multiple, and often very differently measured, markers of complexity, terrorist rhetoric is simpler than ideologically similar but non-terrorist comparison groups.[3]

Affiliation/interpersonal motives

A second clear theme emerging from this dimensional view is that terrorist groups emphasize affiliation and social/interpersonal rhetoric more than non-terrorist groups. In particular, terrorists demonstrate more *rapport* in DICTION, more *social process* words in LIWC, and more *affiliation motives* in motive imagery scoring.

It is worth noting, however, that this focus on things interpersonal seems to be directed primarily at ingroup members. The affiliation motive imagery score was also broken down into affiliation imagery directed towards the ingroup, versus that directed towards the outgroup (this measurement has currently been coded for about 2/3 of the data set; coding of the remainder is ongoing). These results suggest the affiliation difference between terrorist and non-terrorist group rhetoric is entirely accounted for by ingroup affiliation. In the main, it is unlikely that terrorist rhetoric is designed to build rapport with stated terrorist enemies; rather, it is likely to build rapport with sympathizers and potential or current recruits.

Control/power

Another clear theme to emerge is that the terrorist groups' rhetoric is generally higher in measurements of control or power than that of their non-terrorist counterparts. Again, this effect cut across measurements from both automated and manual systems, including significant effects for LTA's *control* measurement; VICS' measurements of *historical control*, the *predictability of the political future*, and the *role of chance*; LIWC's measurement of *status*; and motive imagery's measurement of *power*. (One counter-example was Schwartz' value survey's measurement of *power* – although this effect was on only a portion of the data and was not robust across contexts.) Although these effects were generally more consistent for Control than for Power, in the main they support a larger contention that terrorist rhetoric is, predictably, more focused on gaining power through direct means – and the belief that such power-gaining is, in fact, within their grasp.

Aggression/distrust

One interesting pattern to emerge in the present work is that no significant differences emerged on any variable relevant to aggression/distrust. It is often hard to interpret null findings, but in this case it is noteworthy that the stereotypical view of what distinguishes terrorist from non-terrorist groups – indeed, beyond stereotypes, the actual behavioral dimension that by definition distinguishes them – would suggest that terrorists should show more aggression in their rhetoric. Yet, across every system studied here, the plain fact is that they do not. This further illustrates the sometimes subtle nature of what comprises true terrorist rhetoric: it is not always the seemingly "obvious" things that distinguish terrorists from non-terrorists.

Achievement

Two other themes emerged from this dimensional view, both of which had somewhat conflicting results. First was achievement. Interestingly, there was a strong tendency for achievement-related measures from the automated systems to reveal that terrorist group rhetoric was lower in achievement: *Task orientation* from LTA, *accomplishment* from DICTION, and *achievement* and *work* from LIWC, all showed significantly lower achievement for terrorist rhetoric. However, the most validated manual measurement of achievement – the motive imagery score – showed overwhelmingly the opposite pattern, with terrorists using markedly more achievement imagery. It is hard to know for sure what this means: it is possible that the automated measurements are capturing more of a focus on work and specific tasks, while the motive imagery measure is capturing a deeper drive to succeed writ large. If that is so, it would suggest that terrorists are less likely to talk about the specifics of tasks and task-related terms, but are more likely to implicitly focus on the importance of obtaining success. The achievement language of terrorism may not be easily picked up by word-counting systems. Of course, this is merely speculation.

Optimism/confidence

There were two significant differences for measures related to self-confidence and optimism – differences which pulled in opposite directions. In DICTION, terrorist rhetoric was found to be more optimistic; while in LTA, it was found to be lower in self-confidence. (It is also perhaps worth noting that terrorists exhibited a larger percentage of first-person pronouns, which is relevant to self-confidence or self-promotion.) It is hard to know exactly what this means. One way to reconcile these findings: it is possible that terrorists exhibit more optimism, but not an optimism based in self-confidence – rather their optimism is based in confidence in God, or in the rightness of their cause, or some other external force. This is not totally consonant with the results on belief in control, but is certainly plausible.

Does content analysis give clues to *when* terrorist groups will strike? The terrorist rhetorical style and its consequences

So far, it is clear that these data robustly and consistently define a particular *style* of rhetoric that is uniquely indicative of violent groups. That style is, among other things, more social, more emotion-focused, less complex, and more focused on control/power. Now, we progress to our second (and more complicated) question: can this terrorist style help us understand *when* terrorist groups are going to attack?

Overall analytic strategy for predicting attacks: construction of the warning indices

To measure the likelihood of impending violence, three basic violence-warning variable types were used in the present analyses, each of which has strengths and weaknesses.

Proximity to next attack

First, we used a continuous measurement of *proximity to next attack* (as measured in days to an attack; this measure was inverse-scored so that fewer days equals greater proximity). Higher scores on this measurement meant that the document in question appeared closer in time to an upcoming attack (in other words, proximity here is always forward-looking in time). To the degree that an LCA variable is positively correlated with this proximity variable, that means that as an attack nears, the LCA variable in question became stronger. In contrast, to the degree that an LCA variable is negatively correlated, this means that an impending attack is associated with decreasing levels of the variable in question.

This continuous proximity measurement's primary strengths are that is uses all the documents, and uses the proximity concept to its fullest because it does not force an arbitrary division amongst time frames. As such, in one sense it provides a powerful, broad test of a variable's ability to predict an impending attack. This approach uses all the available data in its "purest" form.

On the downside, this approach also has the potential to miss important effects that occur near the impending attack. Because it treats all documents equally, and because fewer documents were issued especially close to an attack than in other categories, one could easily miss a sudden "drop-off" or "rise" that occurred only in the month before an attack and not at other points. Such a late change might get lost in this large, continuous-measure approach to proximity because it will get "washed out" by the other points in the timeline.

Final month change

As a result of this potential, we also created another variable to capture the likelihood that documents occurring *very* close to an attack were different than the typical document. Documents occurring within one month of an attack by the authoring group were dummy-coded as "1", while all other documents for the group were dummy-coded as "0". Positive correlations between our focal variables and this final month change variable thus mean that there was a tendency for that variable to rise in the final month prior to an attack, compared to the average typical document. Negative correlations mean that there was a tendency for the content analytic variable in question to drop in the final month prior to an attack, compared to the average document. (We present these as correlations for ease of presentation; please note that presenting the means using t-tests would reveal identical p-values.)

This approach has the advantage of getting more at the potential for a "late" change in content analytic variables, and also makes use of all the documents. However, it potentially lowers power because there are a relatively small number of documents in the key ‘‘cell’’ (that is, the one including one month before the attacks). This is less of a problem for al Qa’ida Arabian Peninsula (AQAP), who had a tendency to release more documents in the month prior to an attack. For Central al Qa’ida, who released very few documents (8) in the month prior to an attack by their own group, this approach may yield fewer results.

Focused final month change

One other problem with *both* of the above measures is that they each treat documents that occurred years from an attack as equal in weight to those occurring only months before. It is possible, however, that the best comparison context for materials involves those that appeared close to them in linear time, relative to the attacks. In other words, it may be that the prior analyses might mask a change close to an attack that occurs as a drop relative to the prior month or two, and not a drop relative to things that occurred years before. Since the prior two methods incorporate years of data in each analysis, they may mask this effect. (For example, integrative complexity effects in international crises most frequently occur as relative ‘‘drops’’ or ‘‘rises’’ to the previous statements, not necessarily to some overall baseline established over years; see, e.g. Suedfeld & Bluck, 1988.) To account for this possibility, we created a third measurement, which focused only on the last three months prior to an attack. In this measurement we combined data from 2 to 3 months prior to an attack (dummy-coded as ‘‘0’’) and compared it to data from 1 month prior (dummy coded as ‘‘1’’). Positive correlations between focal variables and this focused final month change variable indicate that, in the last three months, there was a tendency for the last month to show higher scores than the preceding two months; negative correlations mean the focal variable tends to drop in the last month.

Each of these three approaches has complementary strengths and weaknesses; thus, variables that tend to show predictive power across different approaches are more likely to be exhibiting a real relationship with violence.

Central al Qa’ida versus inspired

An additional distinction is worth discussing in the present research. AQAP is somewhat self-contained and thus documents released by that group are counted towards attacks perpetrated by that group, and that group, alone. The situation for Central al Qa’ida (AQ-C) is more complicated. We have in our data set attacks clearly perpetrated by the central group itself, as well as attacks perpetrated by both affiliated groups and inspired groups. How are we to incorporate the affiliated and inspired group attacks?

Perhaps the best way to phrase the issue is in terms of the likelihood of knowledge of the attacks (and help in planning the attacks) from the central group members. One would assume that the person authoring the document for the central group (e.g. bin Laden) would in fact be aware of all attacks attributed to the main group. It is also likely, however, for a transnational organization as large as AQ-C, that they at least had knowledge of affiliated and inspired group attacks – although not as certain as in the case of attacks clearly attributed to the central group.

Thus, we are faced with a classic *number* versus *precision* tradeoff: using the affiliated/inspired groups allows for a larger number of attacks to predict (thus potentially increasing power), but potentially decreases the precision as the certainty of AQ-C foreknowledge and aid decreases. As a result of this tradeoff, we opted to compute two separate measures for AQ-C (for each of the above three approaches to linear time): One measure which used only AQ-C attacks in computation, and one which used AQ-C attacks *plus* attacks attributed to AQ-C-affiliated or AQ-C-inspired groups. This left us with six primary warning indices (3 types of temporal approaches $\times$ two types of group inclusion criteria): (1) Proximity/All, (2) Proximity/Excluding AQ-C-affiliated/inspired, (3) Final Month Change/All, (4) Final Month Change/Excluding AQ-C-affiliated/inspired,(5) Focused Final Month Change/All, (6) Focused Final Month Change/Excluding AQ-C-affiliated/inspired.

Note that for analyses on AQAP separately, we do not need the two distinct group inclusion criteria indices. However, although we computed and present these warning indices for both AQ-C and AQAP separately, for summary purposes we also present these six indices above with both groups combined. These ''combined'' indices all conceptually have the same function: To show the ability of content analytic techniques to predict an impending attack *by the terrorist group that issued the document*. Some of these, for AQ-C, exclude affiliated and inspired group attacks (and are indicated as such), while some of them include attacks by all related groups.[4]

Inclusion/exclusion criteria for documents

It should be noted at that an analysis focusing on predicting specific violent events will necessarily be ''messier'' in multiple ways than a simple across-groups comparison. In addition to an overall reduction in the number of applicable documents (because one cannot meaningfully use non-terrorist documents to predict terrorist activity, the non-terrorist documents were removed for these analyses), terrorists do not release documents in a systematic temporal fashion prior to attacks – sometimes, the closest document is within days of an attack, sometimes it is months from it. Add to this the inherent difficulties of determining the actual date documents were written and the sometimes ambiguous nature of attributing certain attacks to specific groups, and it is easy to see that using content analyses of public documents to predict terrorist activities will inherently involve some degree of messiness. Nevertheless, as we shall see below, there is reason to believe that – although not as clear-cut as distinguishing terrorists from non-terrorists groups – such techniques can in fact help us predict terrorist violence.

Part of this messiness comes from the fact that we have multiple types of documents. Because the document types do not always appear in the same number in each conceptual ''cell'' (e.g. speeches are over-distributed for AQ-C documents close to attacks), and because document types themselves can differ on various linguistic analyses measures irrespective of other variables, this poses a potential confound for any findings. To narrow this possibility, we did two things. (1) First and most importantly, we narrowed the scope of the documents to those three types for which enough information existed to offer legitimate within-document-type comparisons over time: written statements, speeches, and interviews. (2) Second, we also computed the key warning indices for each of the three types separately. While these results predictably yielded occasional differences across document types for specific variables, overall this analysis revealed a remarkably similar pattern across document types. (For example, the key terrorist style summative measures presented below – measures that include a large number of the individual variables from the study – showed similar patterns across the three types of documents.) Indeed, our results encouragingly suggest that future research can successfully make use of multiple types of source materials in understanding terrorist rhetoric and behavior. For ease of presentation, we focus here only on analyses that collapse across document type.

All primary results reported in the tables and below in the narrative are the correlations between each of the six warning indices (two Continuous Proximity measures, two Final Month Change measures, and two Focused Final Month Change measures) and the focal linguistic analyses variables. Across all six indices, positive correlations mean that as proximity to an attack appears, the variable tended to rise, while negative correlations mean the reverse.[5]

A brief dimensional summary

As before, we initially used the same conceptual dimensions presented earlier to gain a birds-eye view of the predictive value of the LTA systems. This approach yielded a mixed set of results that often diverged for automated and human-scored systems. Rather than present a detailed account of each system, we present here only a brief synopsis of these results by conceptual dimension, with the goal of setting up our key terrorist style summary analyses presented below.

For the automated systems, there appeared in the main to be a general tendency for terrorists to exaggerate their typical rhetorical pattern as an attack neared. For example, recall that terrorists were less complex overall than non-terrorists; and, across many of the six indicators of terrorist violence used here, they became even less complex than usual as a violent attack neared. Although occurring in a somewhat inconsistent pattern for many of the other variables, there was a tendency for this same thing to hold for affiliation/interpersonal variables (terrorists generally more interpersonal, became even more so as attacks neared) as well. Overall, although not an overwhelming pattern, for the automated measures, descriptively it appeared that if any predictive value was present, it tended to be that terrorists showed more of their "typical" style immediately prior to an impending attack by their group.

The pattern for the human-scored measures was less clear. For some measures and for some warning indices, this pattern seemed reversed, but overall the human-scored measures showed fairly weak and/or inconsistent predictive value for understanding when terrorist groups would commit violence.

Constructing ad-hoc terrorist style measures

What are we to make of this somewhat inconsistent pattern? To better understand the overall likelihood for terrorists to exaggerate or oppose their "typical" rhetoric as an attack nears, we constructed some ad-hoc measurements from each system here. We call these terrorist style measurements. The conceptual approach was simple: we wanted to construct measurements to represent, in these documents, those specific variables in each system that distinguished terrorist from non-terrorist rhetoric. So we included variables in each system's terrorist style measurement if the variables had shown significant differences in comparing terrorist versus non-terrorist groups.

We focus here on first providing terrorist style scores by system (as opposed to conceptual dimensions). The logic of this is twofold. (1) Each system was designed to be used as a unit, whereas our ad-hoc conceptual dimensions were not. Thus, it makes more methodological sense when trying to define (individual variable by individual variable) a terrorist style, to start *within* each LCA system. (2) This approach – taking each individual variable as used within each system – is methodologically more acute than the broad summaries that focus on dimensions. By focusing *only* on those specific variables that truly defined a terrorist rhetorical style (instead of large conceptual dimensions that include some variables that did not statistically relate to that style), we are able to get better predictive value. (3) Relatedly, this approach allows us to include some variables (e.g. overall "function" words in the LIWC) that, although characteristic of terrorist rhetoric, do not clearly fit in any of the dimensional categories discussed earlier.

What this approach loses in theoretical clarity it thus gains in methodological precision in regards to measuring the over-arching construct of terrorist style. Below, we describe this process of constructing these terrorist style measurements briefly for each system.

Construction of measures

For LIWC, because of the preponderance of variables and their potential overlap, we focused only on larger "summative" categories or measures. Those measures on which terrorists' rhetoric was higher that were included in the final terrorist style measure were: social process, social–emotional style, status, biological terms, positivity, and verbs. On the other side, variables that were included on which terrorist rhetoric was lower were: honesty, complexity, categorical thinking, relativity terms, and overall function words.

Because the conceptual variable we are interested in here is "typical terrorist rhetoric" – and we want higher scores to always represent typical rhetoric – one cannot simply add the positive and negative predictors together (because then extreme positive predictors of terrorist rhetoric and extreme negative predictors would cancel each other out). As a result, we inverse-scored the negative predictors above and averaged them with the positive predictors to get an overall terrorist style score for the LIWC. Higher scores on this measure thus represent an overall style of rhetoric on which terrorists score higher than their non-terrorist counterparts.

We similarly constructed measurements for the other automated systems. For DICTION, we used the 16 primary (non-summative) variables that were predictive of terrorist (versus non-terrorist) rhetoric. Terrorist rhetoric was higher on some of these (self-reference, tenacity, satisfaction, inspiration, human interest, and rapport) and lower on others (numerical terms, collectives, accomplishment, cognitive terms, spatial terms, familiarity, centrality, motion, insistence, and complexity). For LTA, we included the one significant predictor for which terrorist rhetoric scored higher (control), and the two significant predictors for which terrorist rhetoric scored lower (self-confidence and task orientation). For VICS, we included the three measures for which terrorist rhetoric scored higher (flexibility/cooperative-conflict, political future predictable, and historical control) and the two for which terrorist rhetoric scored lower (role of chance and risk orientation). For the LIWC-MEM, we included seven topic domains on which terrorists used significantly more rhetoric, and six for which they used significantly less.

As for the LIWC, for each system we constructed a total terrorist style measure by reverse-scoring the items on which terrorist rhetoric scored lower and averaging the items together. Higher scores indicate, for each system, the more "typical" way that terrorist rhetoric scored on that system. We further averaged these five systems into an overall automated terrorist style measure.

For the manual measures, we used integrative complexity as the summative measurement for that system, and the summative average of all three motive scores for motive imagery. Because terrorist rhetoric typically scored lower on integrative complexity, this variable was reverse scored and the two manual measures were then summed into a total manual terrorist style score, where higher scores represent, for the manual systems, the more "typical" response from a terrorist group member. (Due to a lack of consistent significant findings comparing terrorists and non-terrorist rhetoric – as well as the smaller *n* – we dropped the Schwartz value system from these analyses.)

Key results: does terrorist rhetoric become more "typically terrorist" as an attack nears?

Results from this analysis generally corroborated the more subjective interpretation offered thus far: Namely, that on the automated measures, terrorists showed a marked tendency to exaggerate their "typical" style as an attack neared. Results for all terrorist style measures are presented in Table 7. In particular, it is clear that, for automated measurements (especially looking at warning indices focusing on the last month before an attack), terrorists *do* exaggerate their typical tendencies: For the summative automated terrorist style measure, three of the four late-change warning indices were significantly positive. This suggests that those things that tell us whether or not a particular group is violent writ large, also tell us when a known-violent group is about to engage in violence.

Two additional curious findings from this analysis deserve commentary. First, as Table 7 shows, though the summative automated terrorist style measurement was positively related to 5 of the 6 warning indices, the sixth index (continuous proximity measurement that excludes inspired and affiliated groups) showed a significant pattern in the opposite direction. Why might this be? It may be that focusing only on attacks perpetrated by the focal group yields a kind of curvilinear effect on terrorist rhetoric. It is possible that, under some circumstances, terrorist rhetoric becomes *less* typically "terrorist" in the early planning phases of an attack (thus yielding the negative continuous proximity measure) but then reverse that trend as the attack comes nearer. (And, indeed, curvilinear regression analyses of the proximity–terrorist style relationship show, consistent with this, a significant quadratic effect in this direction.) Of course, it is possible that this pattern is a fluke, as it only occurs for one of the two indices relevant to continuous proximity (with the other index being in the opposite direction).

Second, as shown in Table 7, why might the summative manual terrorist style measurement reveal a pattern somewhat the opposite of the automated measures? There is no clear and easy answer to this question. One possibility is that the automated terrorist style summary score includes not only many more systems, but also *many* more measurements per system. As a result, the automated measurement is more likely to be a reflection of a broad terrorist style, while it is possible that the manual terrorist style measurement merely reflects the idiosyncratic effects of the two

Table 7. Does terrorist style change prior to an impending terrorist attack?

Style variable	Continuous proximity	Continuous proximity excluding aff/insp	Final month change	Final month change excluding aff/insp	Focused Final month change	Focused final month change excluding aff/insp
Terrorist style – LIWC	.02	−.16	.11	.20*	.09	.33*
Terrorist style – DICTION	.26**	−.04	.21*	.16^	.08	.20
Terrorist style – LTA	.20*	−.11	.20*	.18*	.18^	.29*
Terrorist style v VICS	−.07	−.14	.01	.01	.02	.10
Terrorist Style –LIWCMEM	.06	.03	.07	.15^	.07	.21
Terrorist Style –Automated Summary	*.13*	*−.18**	*.19**	*.23***	*.15*	*.40***
Terrorist style – Manual Summary	*−.35****	*−.11*	*−.26****	*−.13*	*−.17*	*−.06*

***$p < .001$; **$p < .01$; *$p < .05$; ^$p < .10$.

measures included in it (integrative complexity and motive imagery). Perhaps if a broader spectrum of manual measurements could be included analogous to the dozens of automated measurements, a similar pattern would emerge for the manual measurements. It is also possible, of course, that the differences reflect substantive differences between automated and human coding writ large, or (more broadly still) that they reflect differences between measurements focused on word counts and those focused on semantic meaning.

General discussion

Summary of key findings

We asked two primary questions in this research. (1) Can linguistic content analysis help us distinguish terrorist groups from ideologically similar non-terrorist groups? (2) Can linguistic content analysis help us understand *when* terrorists are planning to attack? The answer to both questions, with some qualifications, is *yes*. The present results clearly suggest that terrorist rhetoric is robustly different than the rhetoric of groups who share much of their radical ideology: it is (to name but a few examples) more social and emotional, less complex, and more narrative in nature. This work further suggests, albeit more weakly, that one can use this "typical" terrorist style to define a kind of rhetoric that is predictive of *when* terrorists will attack: that, for automated systems at least, terrorists often exaggerate their typical style as an attack nears.

Below, we discuss some qualifications on this research and offer some concluding thoughts.

Qualifications and limitations on the present work

Small number of terrorist groups

One limitation involves the small number of terrorist groups in the study. Although we have talked for ease of presentation of ''terrorist'' rhetoric in describing our results, such language should not be taken to mean that we believe the rhetorical style we are capturing will *universally* apply to all terrorist groups. It will not. Indeed, prior research using one of the systems in this study (the LIWC, which yielded some of the largest and most consistent effects for our project) suggests that different terrorist groups differ in their rhetorical style (Pennebaker & Chung, 2008). Nor would we expect the same predictors of violence to necessarily be in evidence for every group. Indeed, to the degree that (1) terrorist groups tend to exaggerate their rhetorical style as an attack nears and (2) the average rhetorical styles of terrorist groups differ, it logically follows that (3) the specific predictors of terrorist violence will differ somewhat from group to group. It seems unlikely that there will be many (if any) ''one size fits all'' predictors of terrorist violence.

Over and against this, three things should be noted. (1) This is an empirical question on which there are very little data. One should never presume the outcome to a question before the data exist; and thus it is still possible that the results presented here – or at least some of them – represent effects that apply to a broad scope of terrorist rhetoric. (2) The present study examines the rhetoric of two different terrorist groups and multiple sources within each group, and the results we focus on here largely held across both groups. Thus, at the very least, these effects might be thought of as not *just* idiosyncratic to one group only (or at least, indicative of an al Qa'ida style that goes beyond one small group of leaders). (3) Even assuming these effects are idiosyncratic to only the groups here, that does not diminish their worth in predicting violence. Indeed, the effects found here are rather remarkable in their generality – that across multiple automated systems, types of information, and groups, a tendency to exaggerate typical rhetoric emerged during times preceding violence.

The potential for family-wise error

One of the primary issues facing any summary analysis with over 200 variables concerns what statisticians have termed ''family-wise error.'' In particular, we have discussed many significant effects; but given that we had so many linguistic content variables (and, in the case of predicting violence, six outcome variables), is it possible that this just represents the random fluctuation of chance? After all, with 200 variables, it is likely that by chance alone 10 would be statistically significant on any measure.

Given that the effects are so strong, pervasive, and interpretable for the terrorist versus non-terrorist comparisons, this issue is not in evidence there. However, what about for the weaker and less consistent effects of LCA variables on predicting terrorist violence? It is of course (as with any study) possible that these effects are due to chance, and certainly more probable than in a typical, theory-based study with only a handful of variables. (Also note that we did not perform a ''correction'' for family-wise error, which we view as overly ''conservative'' and unnecessary.) A brief look at Table 7 certainly reveals a large percentage of non-significant correlations.

So we frankly and honestly acknowledge this very real possibility. However, two factors at least work against this interpretation. (1) There is a tendency for variables across different systems that are conceptually measuring the same thing to show the same pattern, even when they are not highly correlated with each other. For example, the LIWC complexity, LTA complexity, DICTION complexity, and dialectical complexity measures are not very highly correlated with each other, yet show the same pattern of correlations with violence. The fact that they are *conceptually* related increases the probability that the results are meaningful – because concepts are not related by chance, but by some real underlying theme, and random chance would not likely produce such a pattern. Although this conceptual organization is not perfect by any stretch, even for complexity, there is an overall tendency for certain types of measurements to show the same pattern – especially when one looks at the automated systems. (2) Based on the same logic, and more persuasively, let's turn our attention for a moment to the relationship between terrorist style and violence presented in Table 7. It is worth noting that (a) these terrorist style measurements incorporate a large number of the measurements used in the present study, and (b) they do so in a non-random fashion, organized around an empirically verified, but conceptually coherent, concept. These automated terrorist style measurements, quite remarkably, all show the same pattern, and the terrorist style summary score for the automated measurements shows a clear-cut pattern of predicting change in the final month, where it is significant on three of four measures. (Indeed, as the table on terrorist style shows, 11 of the 30 individual terrorist style measurements are at least marginally significant at $p < .10$; by chance alone, one would expect only 3 to be marginally significant.) So, given the nature of the data, it is unsurprising that the results are messy and sometimes inconsistent; on the other hand, there does seem to be a pattern that emerges here (at least for the automated systems) across a wide array of variables and systems – and such consistent patterns are unlikely to be the result of random chance. If one threw a bucket of confetti in the air, it is *possible* that it may form a heart shape on the ground; it is just not *likely*. Thus, if one *finds* confetti in a heart shape on the ground, however crudely rendered, it is more likely that it reflects not such random forces, but rather the specific intent of a real person.

Limitations in the match between the attack timeline and the documents

It is worth noting that any attempt to categorize variables by "months" is somewhat crude, because they rarely represent an exact sequence by a particular group in each instance. In other words, the documents in the 1-month category did not often occur for the same attack as those in the 2-month category, and so on. As a result of this, there are no attacks for which we have the full sequence of documents (one month prior, two months prior, three months prior, and four months prior) that all occurred *prior to that specific attack*. Terrorists are simply not so cooperative as to release documents in a perfect experimental design. Because we do not have good within-attack comparisons over time, it is possible that our results reflect something other than the implied sequence of (for example) increases in typical terrorist style rhetoric as an attack nears (although it is unclear what that "other" thing might be). However, we suspect that such crudeness would be highly unlikely to *produce* systematically spurious results, but on the contrary would make it more difficult to *find* effects that actually exist (because they basically are adding a lot of random noise). Thus, to the degree that this works against finding an effect, the fact that this data set yields some promising results in spite of this crudeness is impressive.

Comparing automated versus human scoring systems

Although both types of systems were equally able to distinguish group type, in the main it seems that the automated systems produced a more consistent and interpretable pattern of predictive results for the occurrence of violence within terrorist groups than human-scored systems. Why might this be? (1) Part of it is probably that the sheer number of variables produced by the automated systems gives them a higher likelihood of finding legitimate patterns. Indeed, it is doubtful that the strength of the automated systems here occurred because of a greater likelihood of finding a result per measure. Although we did not formally do such an analysis, 1/3 of the violence warning indices for the two summary manual scores that used the entire data set (motive richness and integrative complexity) were significant; and an eyeball test of the automated measures suggests that they did not perform better in this regard. Yet, their strength lies in an ability to produce more variables in a short period of time. While this increases the odds of spurious findings due to family wise error concerns, it also increases the odds of finding legitimate predictors, such as (we believe) the terrorist style measurements in the present work. (2) Part of it may be that the automated systems on average have a higher emphasis on function words – and in this and some other work, those sorts of words are often the best predictors of behavior and psychology. (Indeed, one of the best overall predictors of violence in this data set was the likelihood of using function words writ large.)

Summary thoughts

Taken as a whole, the present project suggests that the use of linguistic content analytic systems can help us understand terrorist violence. These systems were remarkably adept at distinguishing violent from non-violent radical groups, and also appeared useful at predicting *when* violent groups would engage in violence. Both of these findings suggest that terrorists are unintentionally leaving clues, like linguistic fingerprints, to their psychology and behavior in their public rhetoric, and that the systems used here can help us analyze and identify those fingerprints.

Acknowledgments

This material is based upon work supported by the US Department of Homeland Security, Science and Technology Directorate, Human Factors/Behavioral Sciences Division under Inter-Agency Agreement HSHQDC-07-X-00793, made to the Oak Ridge Institute for Science and Education (ORISE). The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the US Department of Homeland Security or ORISE.

Notes

1. In doing so, we largely re-analyze the massive data set described by researchers in this issue, although oftentimes that analysis directly replicates what occurs in the other papers.
2. We also did some analyses to see if document type (e.g. written statement versus speech) moderated any of these effects. These analyses suggested that the effects presented can be viewed as largely independent of the type of document.
3. A counter-example to this might be the two measurements of flexibility from VICS. However, it is worth noting that these measurements are very domain-specific – one

dealing with conflict/cooperation, the other with words/deeds. Thus, although the idea of flexibility is definitely a part of complex thinking – and ''combination'' strategies on those two areas would get higher scores, as judged on that domain, for complexity – those indices are not general measures of each documents' general structural complexity, but only complexity as it pertains to those two issues (when those issues appear).

4. To illustrate: the combined indices look at AQAP documents predicting AQAP attacks, and AQ-C documents predicting AQ-C-related attacks. When combining the indices, the ''All'' indices above contain, for AQAP, the proximity to the next AQAP attack, and for AQ-C, the proximity to the next AQ-C, AQ-C-affiliated, or AQ-C-inspired attack. The ''indices excluding affiliated/inspired group attacks'' contain the same information for AQAP as the ''All'' indices, but for AQ-C contain only attacks attributed to the central group.
5. Sometimes the results we report here are inconsistent with those in the accompanying individual reports written by the senior researchers. When this occurs, it is almost always the result of different exclusion criteria (e.g. including/excluding book chapters and other materials, including/excluding materials by group members that were not leaders). A discussion of each discrepancy is beyond the scope of this report. The analyses presented here summarize a lot of data in a consistent manner; see the individual reports for a more narrow focus on specific issues.

References

Pennebaker, J.W. & Chung, C.K. (2008). Computerized text analysis of al-Qaeda statements. In K. Krippendorff & M. Bock (Eds.), *A content analysis reader* (pp. 453–466). Thousand Oaks, CA: Sage.

Smith, A.G., Sudefeld, P., Conway, L.G. III., & Winter, D.G. (2008). The language of violence: Distinguishing terrorist from nonterrorist groups by thematic content analysis. *Dynamics of Assymetric Conflict*, *1*, 142–163.

Suedfeld, P., & Bluck, S. (1988). Changes in integrative complexity prior to surprise attacks. *Journal of Conflict Resolution*, *32*, 626–635.

Suedfeld, P., & Leighton, D.C. (2002). Early communications in the war against terrorism: An integrative complexity analysis. *Political Psychology*, *23*, 585–599.