Psycholinguistic Signals of Terrorist Attacks

Natasha Mather (nmather@albany.edu) and Michael D. Young (myoung4@albany.edu)
University at Abany
WORKING PAPER PLEASE DO NOT QUOTE

One of the many problems associated with terrorism is the unpredictable timing of terrorist attacks which makes protecting both the public and critical infrastructure difficult because identifying predictive variables is an active area of terrorism research (Munk, 2017; Meng, Nie, & Song, 2019; Huamaní, Alva, & Roman-Gonzalez, 2020; Abdalsalam, Li, Dahou, & Noor, 2021). Although few studies predict the timing of attacks, it remains a worthwhile goal. Despite some notable efforts (Lazarevska et al, 2005; Smith et al, 2008; Ebner, Kavanagh, Whitehouse, 2023), analyzing the rhetoric of individuals and/or groups has been neglected in favor of "big data" approaches. The analysis of rhetoric allows analysts to explore changes in how a person communicates (verbs, emotions, motivations), rather than just what a person says. Researchers (Hogenraad and Garagozov, 2008) use rhetoric analysis programs to analyze world leaders, exploring changes in rhetoric before a decision is made. In some cases, researchers hoped to have forewarning before the decision to use violence is enacted by said group or individual (Hermann & Sakiev, 2011; Walker 2011; Winter, 2011, Conway III, Gornick, Houck, Towgood & Conway, 2011). For example, Hogenraad and Garagozov (2008) looked at documents from the Russian President and Georgian President before August 2008 to scale the risk of conflict by analyzing changes in rhetoric. They found that there were changes in rhetoric before violence broke out in Georgia in August of 2008. Not only has research been conducted on the rhetoric of political figures, but also terrorists/terror groups.

Over the years, several researchers have explored how to use rhetoric to determine new ways to analyze terror groups (Conway and Conway, 2011; Smith, 2008; Smith, 2004; Smith, 2011). One of the primary goals of this research was to determine if there could be forewarning before a terrorist attack took place. To determine this, researchers measured different psycholinguistic variables in rhetoric and various times before an attack. For example, Smith explored if motive imagery (Need for Power and Need for Affiliation) could differentiate terrorist groups from non-terrorist groups with similar ideologies (Smith, 2008). Smith concluded that there was a difference in motive imagery between the terrorist and non-terrorist groups. Terrorist groups scored higher in power and affiliation motive imagery than the non-terrorist groups. In 2011, Winter conducted a similar study to determine if motive imagery could differentiate terrorist from non-terrorist groups and/or give warning to an upcoming attack. He was able to confirm that terrorist groups score higher than non-terrorist groups in power and affiliation motive imagery. There was also a difference in scores between the two tested terrorist groups, AQ-C and AQAP (Winter, 2011). However, Winter found little evidence that any of the motive imagery variables could predict terrorist attacks.

Although there have been studies that show a correlation between an increase/decrease of a psycholinguistic variable and an impending attack (Pennebaker, 2011; Hermann & Sakiev, 2011), no study has been able to predict an attack within a short time limit. Rather than attempt to directly develop models to predict the timing of terrorist attacks, our initial research seeks to determine if the scores of psycholinguistic variables have an impact on the length of the interval between terrorist attacks. Survival analysis is the best approach to testing the effect of the psycholinguistic variables on the interval between terror attacks. In this model, the presence of a particular psycholinguistic score is analogous to the administration of a treatment. A shorter time

frame would represent an increase in terrorism for a group. Hence, in a shorter time frame we would expect to see a change in score for the psycholinguistic variables. To explore the impact psycholinguistic variables, have on the length of the interval between terror attacks, several psycholinguistic variables previously associated with violence are assessed:

- Need for Power (nPWR)
- Need for Achievement (nACH)
- Belief in the Ability to Control Events (BACE)
- Distrust in Others (DIS)
- View of the Political Universe (P1)

Independent variables Tested:

Need for Power (nPWR)

The need for power variable (Winter, 2011) measures a leader's desire to establish, maintain, and/or restore power and authority over others. In the documents, this variable is coded using Young's (forthcoming) automated coding scheme for Profiler Plus (Levine & Young, 2014) in the motive system when an individual uses verbs that relate to desiring power over others. The nPWR score for a document is determined by counting the number of times a verb occurs in the document, converting that image score per 100 words. An increase in nPWR is positively associated with a higher probability of violent action (Smith, 2008). In terms of terrorism, as the proximity of a terrorist attack nears, the need for power is expected to increase (Smith, Suefeld, Conway III, Winter, 2008; Suedfeld, Brcic, 2011; Winter, 2011).

Need for Achievement (nACH)

The need for achievement variable measures an individual or group's desire to perform at a standard of excellence and/or achieve goals (Winter, 2011). In the documents, researchers look for positive adjectives and verbs that mention unique accomplishments, winning, performance, and fear of failure (Winter, 2011). The need for achievement is also coded in the motive system, which takes raw scores for the variable and converts it into a motive score per 1000 words. In previous studies, nACH is positively associated with a higher probability of violent action. As a terrorist attack nears, nACH increases (Lazarevska, Sholl, Young, 2005; Hermann & Sakiev, 2011).

Belief in the Ability to Control Events (BACE)

BACE is an LTA variable that measures how individuals/groups view their ability to control events that will or are currently happening around them. When coding documents, the variable is scored when the speaker has taken responsibility for enacting/planning an action. A higher BACE score is associated with an increased likelihood of engaging in terrorism (Smith, Suefeld, Conway III, Winter, 2008; Suedfeld, Brcic, 2011).

Distrust in Others (DIS)

DIS is a Leadership Trait Analysis (LTA) variable that measures the feelings of guilt, mistrust, and weariness an individual or group places on others (Hermann & Sakiev, 2011). In the documents, nouns and noun phrases are coded when the individual voices concern or mistrust of others. A higher DIS score is associated with an increased likelihood of engaging in terrorism (Lazarevska, Sholl, Young, 2005; Hermann & Sakiev, 2011).

View of the Political Universe (P1)

P1 is an operation code variable that measures how an individual views themself in the political world around them (Walker, 2011). This explores if they find the world cooperative or hostile, whether the behavior of others is predictable or not, and the way others can interfere with the individual's power (Walker, 2011). In the documents, verbs are coded based on interpretation, a +1 indicated friendly and a –1 indicated pessimistic. In the existing literature (Walker, 2011), P1 is negatively associated with terrorism.

From this literature several hypotheses can be derived:

H_1: BACE is positively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005; Hermann and Sakiev, 2011)

H_2: nPOW is positively associated with a shorter time interval between terror attacks (Smith, Suefeld, Conway III, Winter, 2008; Suedfeld, Brcic, 2011; Winter, 2011)

H_3: nACH is positively associated with a shorter time interval between terror attacks (Smith, Suefeld, Conway III, Winter, 2008; Suedfeld, Brcic, 2011)

H_4: DIS is positively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005; Hermann & Sakiev, 2011)

H_5: P1 is negatively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005; Walker, 2011)

Research Design.

To determine if there was a link between changes in terrorist linguistic rhetoric before an attack, we used survival analysis. Survival (time to event) analysis is a statistical model that uses cumulative incidence to analyze the expected period before an event(s) happens. This allows us

to evaluate the effect of the presence of specific speech on the length of time between terrorist attacks measured by the number of days in the interval. This model is used frequently in medical studies, as it can measure the efficacy of drugs or vaccines (a treatment). In our analysis we used this statistical model to determine if the presence of high or low scores for rhetoric changed the interval between attacks. To assess the impact that rhetoric can have in predicting terrorist attacks, we coded each time interval between attacks to separately determine the presence or absence of communications with high scores for BACE, nPWR, nACH, and DIS, and low scores for P1.

Commented [YD1]: need to pick one way of describing these variables: rhetoric/linguistic content/psychological variables.

This study focuses on three terror groups: Hamas, Hezbollah, and Ansar Allah (Houthis). All three terrorist groups are unique as they are also present in government positions within their countries. Hamas and the Houthis are the main governing fore within the Gaza and Yemen. Hezbollah occupies positions within the Lebanese government and has its own standing army which is allowed to operate in Lebanon. Hamas, Hezbollah, and the Houthis are allied with one another and share similar state sponsors, namely Iran and Syria. To analyze the impact rhetoric has on signaling an upcoming terrorist attack from the groups, documents were collected from leaders, generals, and spokespersons of the groups.

Commented [YD2]: Add a list of the variables that will be used with brief explanations: For example: length of interval between attacks.

Hamas

Hamas is a Muslim Palestinian terrorist organization that was founded in 1988 as an offshoot of the Muslim Brotherhood. The newly formed group played an essential role in the First Intifada, educating youth against Israel and forming a social order in areas it controlled. A power struggle began to form between Hamas and Fatah, the Palestinian group led by Arafat. In 2000, the second intifada would begin, from 2000-2002, Hamas claimed responsibility for 40 suicide attacks, killing over 400 Israelis. Hamas and the PLO would continue their power struggle

against one another, in 2007 Hamas won control over Gaza where they have remained in power since. The group has strong ties with Iran, as they receive monetary, weaponry, ideological, and training support from the Iranians. In 1993, Iran gave Hamas 15 million dollars to fight the peace negotiations between Israel and the PLO. During the heightened conflict between Israel and Gaza, Iran provided weapons and training to the group. In May of 2021, during the conflict between Hamas and Israel, Hamas leader Ismail Haniyeh said that Iran "did not hold back with money, weapons, and technical support."

Hezbollah

Hezbollah is a Shiite Lebanese terrorist organization that was formed in the early 1980s after the Islamic Revolution in Iran (1979) and the Israeli invasion of Lebanon (1982). The Lebanese terrorist group received funding and ideological support from Iran, its main state supporter. Throughout Lebanon's civil wars and conflict with neighboring countries, Hezbollah has been a main political and military force. Hezbollah is represented in the Lebanese parliament and is allowed its military force in the country. The terrorist group is idolized by other Shiite extremist groups in the region, with groups in Iraq, Bahrain, Saudi Arabia, and Afghanistan naming themselves after Hezbollah and attempting to follow a similar path. In recent years, Hezbollah has become directly involved in the Syrian civil war, fighting alongside Bashar al-Assads SAA forces and other Iranian-backed fighters to keep the Syrian president in power. Hezbollah's relationship with Iran continues to grow, as Iran sends support in terms of money, equipment, weapons, and training to the terrorist group.

Ansar Allah (Houthis)

Ansar Allah, also known as the Houthis, is a Zaydi Shiite terrorist group that is in northern Yemen. The terrorist group was founded in the 1990s to fight against the Yemeni government. The uprising failed and in 2004 the leader, Hussein al-Houthi, was captured and executed by the Yemeni government. In 2011, the Arab Spring erupted, and protests began against Yemeni President Saleh, the Houthis played a vital role in overthrowing the president. When a new government, led by President Hadi, took over, The Houthis staged a coup and took over the capital, Sanaa, in 2014. This prompted a devastating civil war in an already volatile country, with actors like Saleh, Hadi, Al-Qaeda, Houthis, southern separatists, and tribesmen fighting amongst each other to gain power. The conflict became even more complex when in 2015, Saudi Arabia and its allies began an intervention in Yemen to fight against the Houthis. Since 2015, the Iranian-Houthi relationship continues to get stronger, with Iran sending weapons (which it denies) to the Houthi militants. Years of conflict in Yemen, alongside strong support from Iran which has allowed the Houthi militants to block a Saudi advance, has given the terrorist group credibility and experience to be a credible threat to the US and its Arab Gulf allies in the region.

Data Collection

Documentation was collected from three sources of communication, these included written press releases, translated video recordings of the terrorist leaders, translations of interviews of terrorist leaders, and tweets from verified generals and/or leaders. Memri (Middle East Media Research Institute) TV was used for all three terrorist groups. The institute collects video and article sources from different Middle Eastern news agencies and translates clips into English. Traditional news outlets like CNN, BBC, Al-Jazeera, and Reuters were also used when they interviewed terrorist leaders or when they included quotes from leaders.

For documents on the Hezbollah leader Hassan Nasrallah, Al-Manar TV was used, they are owned and operated by Hezbollah. Al-Manar has a website in English where Nasrallah's speeches are translated into English. For the Houthis, the Yemeni news agency Al-Masirah was used to collect documents for the group, this site is controlled by the militant group. Al-Masirah has both an Arabic and English language website, their English website contains statements and translated videos of the terrorist leaders. Houthi brigadier general, Yayha Saree, has a verified Twitter account where most tweets are in English. On days when Saree tweeted more than 4 sentences of information, the tweet data was collected and implemented into a Notepad file. The account also contains speech videos and press releases from the Houthis that have been translated into English. For Hamas, Memri TV was the primary source of documentation, as well as sources like BBC and other traditional news outlets that interviewed leaders. The videos appearing on Memri TV came from Palestinian media like al-Aqsa TV, which is owned and controlled by Hamas.

Documentation for Hezbollah leader, Hassan Nasrallah, was plentiful and had credible translations. Most documents also contained more than a few lines of text, with some containing more than 50 lines. The documents were also inclusive of all time periods, with Nasrallah releasing speeches, press releases and/or interviews weekly. For Hamas, most of the documents came from Memri TV, this provides great credibility in their translations. However, Memri only posts clips of terrorist speeches, so the documents used for Hamas typically only consisted of 5-10 lines of data. Since there were 6 leaders that represented Hamas, all time periods of the group were covered in the documents. The Houthis presented the biggest challenge for collecting documentation.

Out of the three terrorist groups, the Houthi movement receives the smallest amount of international reporting. Their English website only contained videos and press releases going back to early 2018. For Memri TV, most of their reporting was after 2018 as well, with only three videos taking place prior. Hence, for this research study documentation was only collected from 2018 onwards, unlike the other two groups and Iran who go back to 2013. The reporting from al-Masirah had credible translation, but their documentation only contained a few lines from every speech, this left Houthi documentation smaller than the other groups. If another study is conducted using the Houthis, an Arabic translator should be used to translate Houthi documentation. On the internet, more documentation can be collected for the Houthis but is only presented in Arabic.

By using survival analysis in NCSS, we were able to test our hypothesis on the group variables. If the p-value was $p < 0.05$, we accepted the alternate hypothesis. If the p-value was $p > 0.05$, we accepted the null hypothesis. There are a number of statistical tests for significant differences between survival curves. As we were unable to determine the most appropriate significance test in this context, we treat each case (group* variable) as significant if at least two of the three common tests have p-values ≤ 0.05 (Gray's, Pepe and Mori, and log-rank).

We tested survival curves for two groups of intervals: 1) intervals with speech scoring above the mean and 2) intervals without speech or speech scoring below the mean and 2). We repeat this process for 1] intervals with speech more than one standard deviation above the mean and 2] intervals without speech or speech scoring below the mean plus one standard deviation. These thresholds reflect a simple definition of 'high' above the mean and a definition of 'high' as one standard deviation above the mean that follows common usage in the literature (Hermann,

Commented [MK3]: selection criteria.

2011). This division is altered for P1 where low values are expected to matter rather than high. The following will go over each of the 5 hypothesis and explore their results:

BACE (Belief in Ability to Control Events)

H_0: BACE is not associated with a shorter time interval between terror attacks

H_1: BACE is positively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005; Hermann and Sakiev, 2011)

Table 1. Difference in Survival Curve Results for Belief in Ability to Control Events (BACE)

Group	Threshold Value	P-Value	Reject H_0	Interval
Hamas	0.368131	**0.00004** (Gray's)	yes	longer
		0.00000 (Pepe and Mori's)		
		0.00000 (Logrank)		
	0.644455	**0.04251** (Gray's)	yes	longer
		0.01405 (Pepe and Mori's)		
		0.00122 (Logrank)		
Hezbollah	0.368131	**0.00923** (Gray's)	yes	longer
		0.00098 (Pepe and Mori's)		
		0.00165 (Logrank)		
	0.644455	**0.01913** (Gray's)	yes	longer
		0.00240 (Pepe and Mori's)		
		0.00487 (Logrank)		
Houthis	0.368131	0.17919 (Gray's)	yes	crossover
		0.00007 (Pepe and Mori's)		
		0.04251 (Logrank)		
	0.644455	0.49790 (Gray's)	no	crossover
		0.06251 (Pepe and Mori's)		
		0.39389 (Logrank)		

Hamas results: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, all three tests have p-values ≤ 0.00004 (Table 1) therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher BACE score is associated with a longer interval between terrorist attacks (Figure 1). When using one standard deviation above the mean as the threshold all three tests have p-values ≤ 0.04251 (Table 1) and was also associated with a longer time interval between terrorist attacks (Figure 1).

Commented [MK4]: check this.

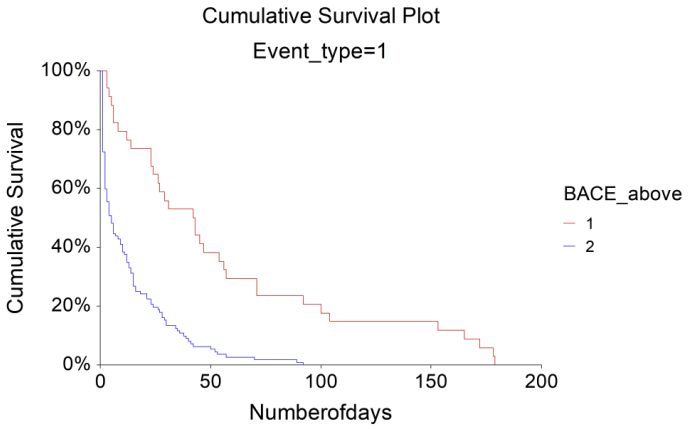

Figure 1. Cumulative Survival Plot for Hamas for BACE above the mean.

Hezbollah result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the p-value ranges from .0092-.0191 (Table 1). Since both p-values scores are ≤.05, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher BACE score was associated with a longer time interval

between terrorist attacks (Figure 2). Furthermore, one standard deviation above the mean had a p-value between .0049 and .0191 (Table 1) and was also associated with a longer gap between terrorist attacks (Figure 2).

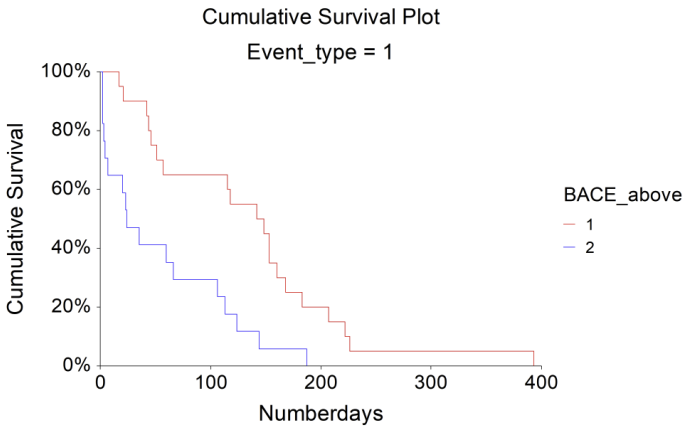

Figure 2. Cumulative Survival Plot for Hezbollah for BACE above the mean.

Houthis (Ansar Allah) result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the p-value ranges from .00007 to .17919 (Table 1). Since two out of the three tests show significance, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher BACE score was associated with a longer time interval (Figure 3) between terrorist attacks. One standard deviation above the mean had a p-values between .06251 and .49790, hence these values were not significant (Table 1).

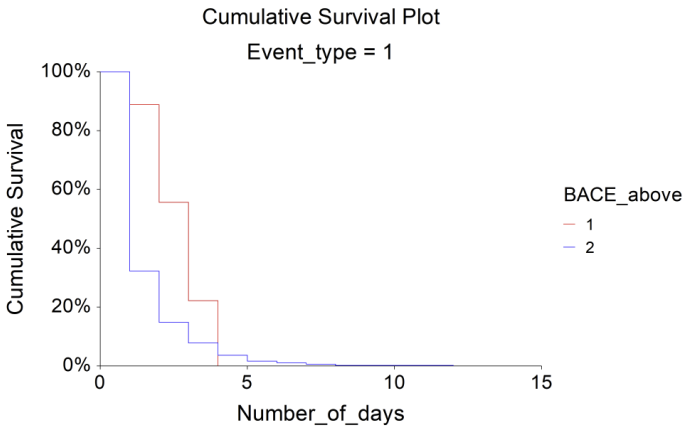

Figure 3. Cumulative Survival Plot for Houthis for BACE above the mean.

Need for Power

H_0: nPOW is not associated with a shorter time interval between terror attacks

H_2: nPOW is positively associated with a shorter time interval between terror attacks (Smith, Suefelf, Conway III, Winter, 2008; Suedfeld, Brcic, 2011; Winter, 2011).

Table 2. Difference in Survival Curve Results for Need for Power (nPWR)

Group	Threshold Value	P-Value	Reject H_0	Interval
Hamas	0.036919	**0.00022** (Gray's)	yes	longer
		0.00000 (Pepe and Mori's)		
		0.00000 (Logrank)		
	0.059658	0.32295 (Gray's)	no	longer
		0.12154 (Pepe and Mori's)		
		0.14200 (Logrank)		
		0.00335 (Gray's)		

Hezbollah	0.036919	**0.00001** (Pepe and Mori's) **0.00041** (Logrank)	yes	longer
	0.059658	0.35705 (Gray's) **0.00010** (Pepe and Mori's) 0.24734 (Logrank)	no	crossover
Houthis	0.036919	0.06221 (Gray's) **0.00000** (Pepe and Mori's) **0.0008**1 (Logrank)	yes	crossover
	0.059658	0.13275 (Gray's) **0.00001** (Pepe and Mori's) **0.00685** (Logrank)	yes	crossover

Hamas result: Dividing the intervals without speech or those with speech above and below the mean, the p-values range from .0000 to .0002 (Table 2), hence the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher mean nPWR score was associated with a longer time interval between terrorist attacks (Figure 4). Interestingly, when one standard deviation above the mean was tested the p-values ranged from .12 to .39 (Table 2). Hence, for Hamas one standard deviation above the mean is not significant. One standard deviation above the mean also had a longer time interval between terror attacks (Figure 4).

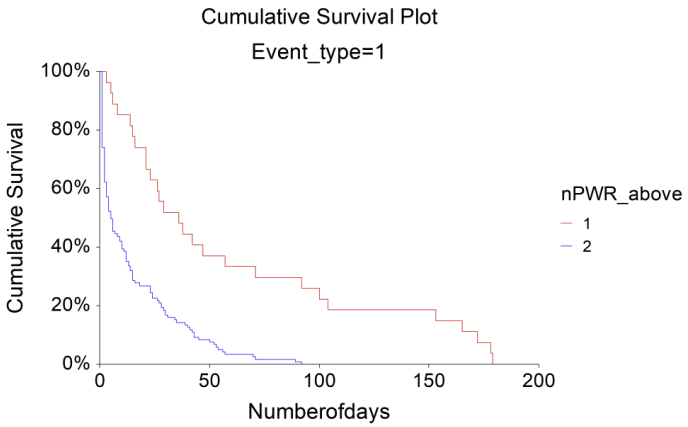

Figure 4. Cumulative Survival Plot for Hamas for nPWR above the mean.

Hezbollah result: Dividing the intervals without speech or those with speech above and below the mean, the p-values range from .0000 to .0034, hence the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected (Table 2). As shown in the results below in the nPWR table, a higher mean nPWR score was associated with a longer gap between terrorist attacks (Figure 5). Like Hamas, one standard deviation above the mean had p-values that ranged from .0001 to .3571 (Table 2). Since two (Gray's and log-rank) out of three of the tests were not significant, one standard deviation above the mean for nPWR is not significant for Hezbollah. One standard deviation above the mean also had a longer time interval between terror attacks (Figure 5).

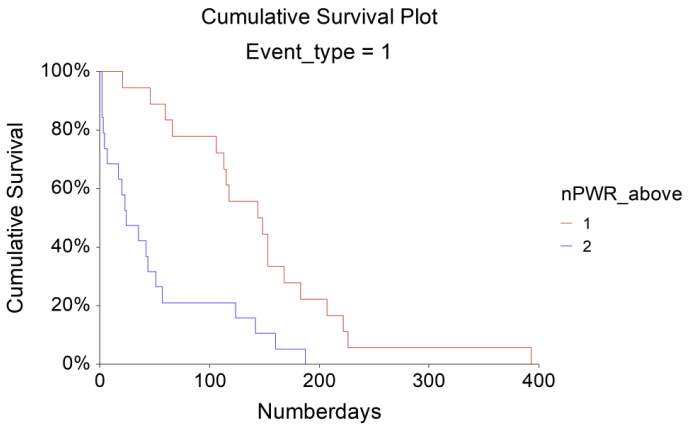

Figure 5. Cumulative Survival Plot for Hezbollah for nPWR above the mean.

Houthis (Ansar Allah) result: Dividing the intervals without speech or those with speech above and below the mean, the p-values range from .0000 to .06221 (Table 2). Since two out of the three tests show significance, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher BACE score was associated with a longer time interval between terrorist attacks (Figure 6). One standard deviation above the mean produced similar results, with p-values ranging from .00001 to .13275 (Table 2). Since two (Pepe and Mori's and log-rank) out of the three tests were significant, one standard deviation above the mean for nPWR is significant for the Houthis.

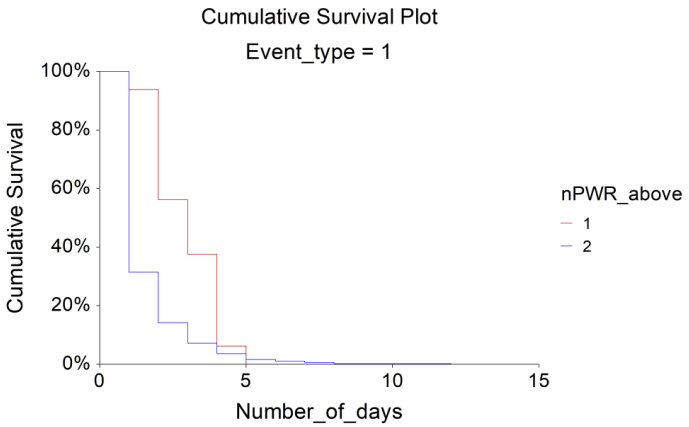

Figure 6. Cumulative Survival Plot for Houthis for nPWR above the mean.

Need for Achievement

H_0: nACH is not associated with a shorter time interval between terror attacks.

H_3: nACH is positively associated with a shorter time interval between terror attacks (Smith, Suefelf, Conway III, Winter, 2008; Suedfeld, Brcic, 2011)

Table 3. Difference in Survival Curve Results for Need for Achievement (nACH)

Group	Threshold Value	P-Value	Reject H_0	Interval
Hamas	0.005573	**0.03926** (Gray's)	yes	longer
		0.00266 (Pepe and Mori's)		
		0.00016 (Logrank)		
	0.012338	0.68926 (Gray's)	no	longer
		0.44642 (Pepe and Mori's)		
		0.63620 (Logrank)		
		0.35541 (Gray's)		

Hezbollah	0.005573	0.09365 (Pepe and Mori's)	no	cross
		0.29659 (Logrank)		
	0.012338	0.36550 (Gray's)	no	cross
		0.31369 (Pepe and Mori's)		
		0.25146 (Logrank)		
Houthis	0.005573	0.22045 (Gray's)	yes	crossover
		0.00476 (Pepe and Mori's)		
		0.00239 (Logrank)		
	0.012338	0.30808 (Gray's)	no	crossover
		0.16140 (Pepe and Mori's)		
		0.02899 (Logrank)		

Hamas result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the three tests have p-values that range from .0393 to .0002 (Table 3). Therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher mean nACH score was associated with a longer time interval between terrorist attacks (Figure 7). One standard deviation above the mean was not significant, as the p-value scores were ≥.44642 (Table 3).

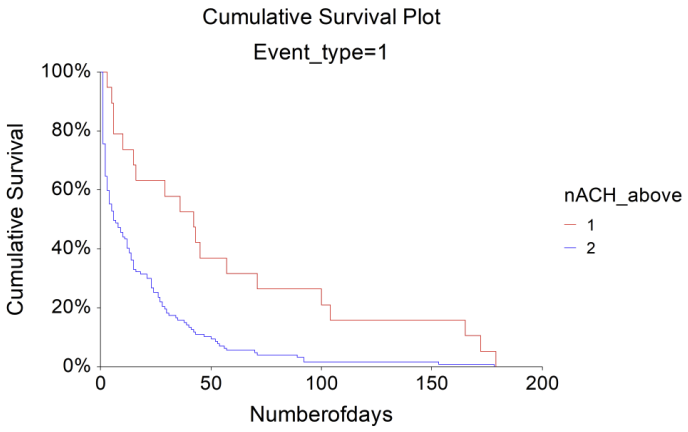

Figure 7. Cumulative Survival Plot for Hamas for nACH above the mean.

Hezbollah result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the three tests have p-values from .0937 to .3554 (Table 3). Hence, since all the p-value scores were above .05, the null hypothesis was accepted. One standard deviation above the mean yielded similar results, with the p-values ranging from .2517 to .3656 (Table 3). Hence, one standard deviation above the mean was also not significant.

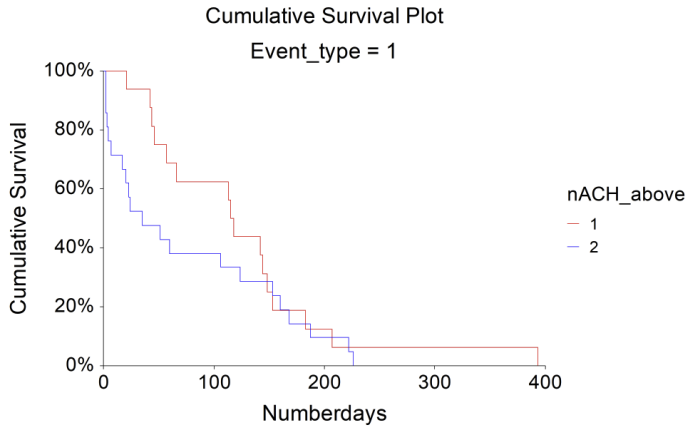

Figure 8. Cumulative Survival Plot for Hezbollah for nACH above the mean.

Houthis (Ansar Allah) result: Dividing the intervals without speech or those with speech above and below the mean, the p-values range from .00239 to .22045 (Table 3). Since two out of the three tests show significance, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. However, one standard deviation above the mean was not significant, as two (Grey's and Pepe and Mori's) out of the three tests were not significant (Table 3).

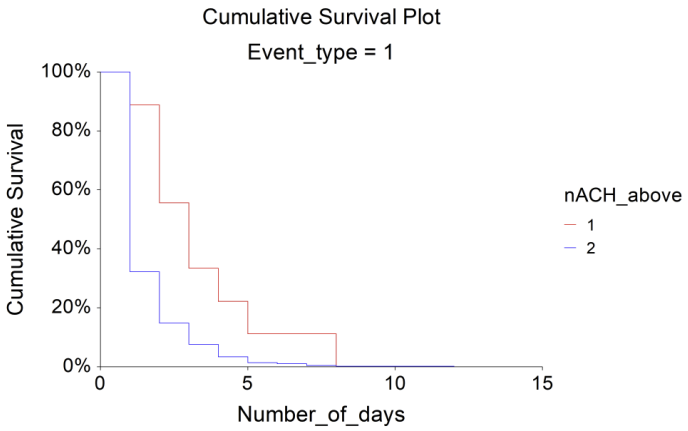

Figure 9. Cumulative Survival Plot for Houthis for nACH above the mean.

DIS (Distrust in Others)

H_0: DIS is not associated with a shorter time interval between terror attacks.

H_4: DIS is positively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005)

Table 4. Difference in Survival Curve Results for Distrust of Others (DIS)

Group	Threshold Value	P-Value	Reject H_0	Interval
Hamas	0.328989	**0.02894** (Gray's)	yes	longer
		0.00023 (Pepe and Mori's)		
		0.00003 (Logrank)		
	0.587576	0.18693 (Gray's)	no	crossover
		0.04299 (Pepe and Mori's)		
		0.06611 (Logrank)		
		0.02594 (Gray's)		

Hezbollah	0.328989	**0.01057** (Pepe and Mori's)	yes	longer
		0.01097 (Logrank)		
	0.587576	0.82234 (Gray's)	no	crossover
		0.08057 (Pepe and Mori's)		
		0.80536 (Logrank)		
Houthis	0.328989	0.07425 (Gray's)	yes	crossover
		0.00000 (Pepe and Mori's)		
		0.00057 (Logrank)		
	0.587576	0.24934 (Gray's)	yes	crossover
		0.00114 (Pepe and Mori's)		
		0.02856 (Logrank)		

Hamas result: Dividing the intervals without speech or those with speech above and below the mean, the p-values from the three tests ranged from .0000 to .02894 (Table 4). Therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A higher DIS score is associated with a longer time interval between terrorist attacks (Figure 10). When using one standard deviation above the mean as the threshold, two (Gray's and log-rank) out of three tests have a p-value over .05 (Table 4). Therefore, one standard deviation above the mean is not significant.

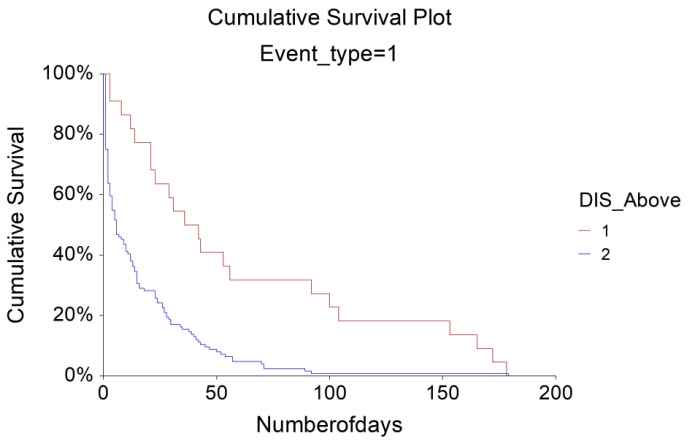

Figure 10. Cumulative Survival Plot for Hamas for DIS above the mean.

Hezbollah result: Dividing the intervals without speech or those with speech above and below the mean, the p-values from the three tests ranged from .0260 to .0110 (Table 4), hence the null hypothesis (H_0) and the alternative hypothesis (H_1) was rejected. A higher mean DIS score was associated with a longer time interval between terrorist attacks (Figure 11). All the p-values from the three tests for one standard deviation above the mean was above .05 (Table 4), hence one standard deviation above the mean is not significant.

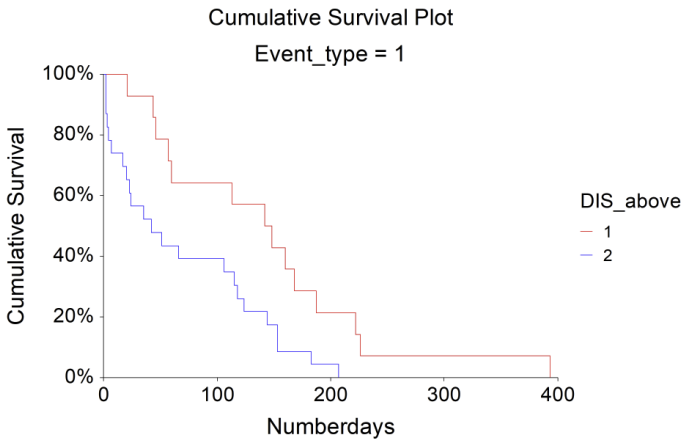

Figure 11. Cumulative Survival Plot for Hezbollah for DIS above the mean.

Houthis (Ansar Allah) result: Dividing the intervals without speech or those with speech above and below the mean, the p-values for two out of the three tests were significant (.00114 and .02856) (Table 4), hence the null hypothesis (H_0) and the alternative hypothesis (H_1) was rejected. A higher mean DIS score was associated with a longer time interval between terrorist attacks (Figure 12). Similarly, two out of the three tests were significant for one standard deviation above the mean (Table 4).

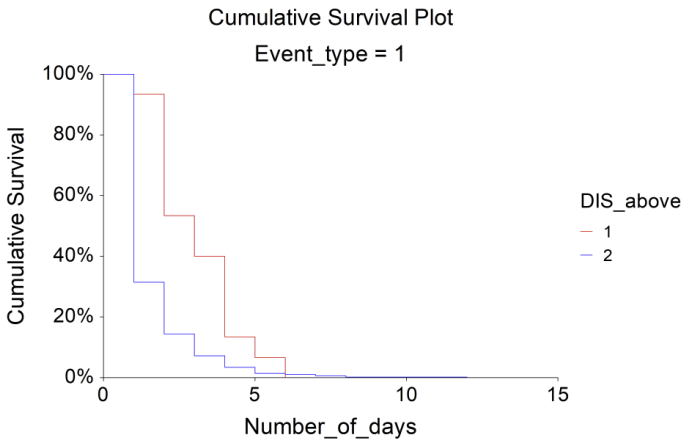

Figure 12. Cumulative Survival Plot for Houthis for DIS above the mean.

P1 (Political Universe)

H_0: P1 is not associated with a shorter time interval between terror attacks

H_5: P1 is negatively associated with a shorter time interval between terror attacks (Lazarevska, Sholl, Young, 2005)

Table 5. Difference in Survival Curve Results for Political Universe (P1)

Group	Threshold Value	P-Value	Reject H_0	Interval
Hamas	-0.03525	**0.03518** (Gray's)	yes	longer
		0.00018 (Pepe and Mori's)		
		0.00006 (Logrank)		
	-0.58454	0.14647 (Gray's)	yes	longer
		0.02147 (Pepe and Mori's)		

		0.00837 (Logrank)		
Hezbollah	-0.03525	**0.00561** (Gray's)	yes	longer
		0.00028 (Pepe and Mori's)		
		0.00195 (Logrank)		
	-0.58454	**0.01789** (Gray's)	yes	longer
		0.00205 (Pepe and Mori's)		
		0.00428 (Logrank)		
Houthis	-0.03525	0.08317 (Gray's)	yes	crossover
		0.00000 (Pepe and Mori's)		
		0.00089 (Logrank)		
	-0.58454	0.08317 (Gray's)	yes	crossover
		0.00000 (Pepe and Mori's)		
		0.00089 (Logrank)		

Hamas result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the three tests have p-values .03518 or below (Table 5). Therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A lower mean P1 score was associated with a longer time interval between terrorist attacks (Figure 13). Furthermore, when testing for one standard deviation below the mean, two (Pepe and Mori's and log-rank) out of the three tests were statistically significant (Table 5). One standard deviation above the mean was also associated with a longer time interval between terrorist attacks (Figure 13).

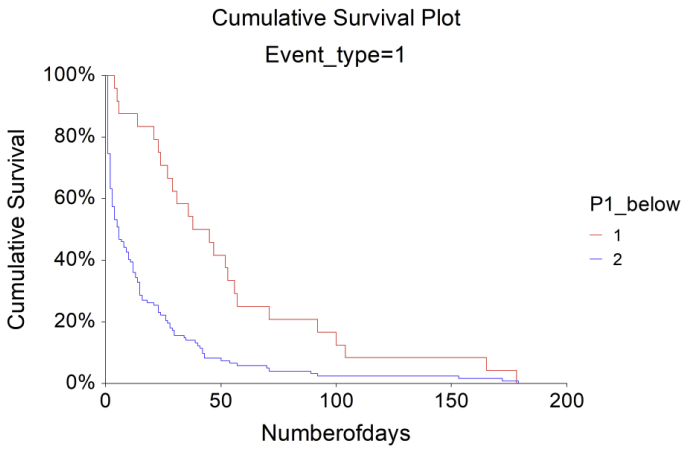

Figure 13. Cumulative Survival Plot for Hamas for P1 below the mean.

Hezbollah result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, the three tests have p-values .00561 or below (Table 5). Therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. A lower mean P1 score was associated with a longer time interval between terrorist attacks (Figure 14). Furthermore, when testing for one standard deviation below the mean, all three tests were statistically significant (≤00195). One standard deviation above the mean was also associated with a longer time interval between terrorist attacks (Figure 14).

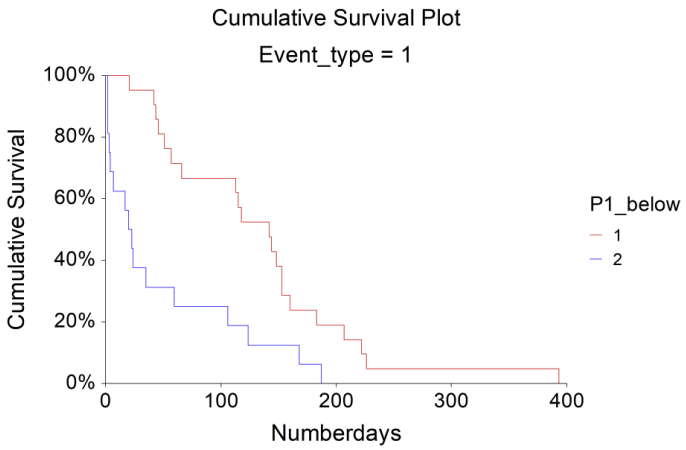

Figure 14. Cumulative Survival Plot for Hezbollah for P1 below the mean.

Houthis (Ansar Allah) result: When the intervals are divided between those without speech or speech scoring below the mean and those with speech above the mean, two (Pepe and Mori's and log-rank) out of the three tests came back significant (Table 5). Therefore, the null hypothesis (H_0) and the alternative hypothesis (H_1) were rejected. Similarly, when testing one standard deviation below the mean, two (Pepe and Mori's and log-rank) out of the three tests came back significant (Table 5).

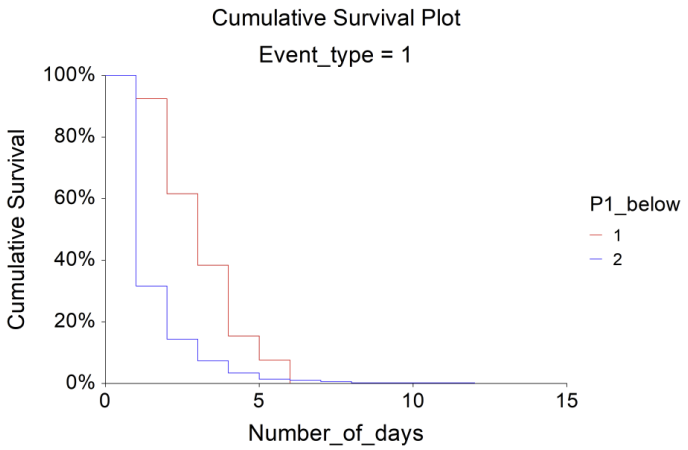

Figure 15. Cumulative Survival Plot for Houthis for P1 below the mean.

A Deeper Dive

The results we have generated do not match up with our predicted hypotheses. Looking at previous studies on BACE, nPWR, nACH, and DIS, we believed that higher scores would lead to shorter intervals between attacks. Similarly, for P1, our prediction was that a negative score for this variable would lead to a shorter interval between attacks, but this was not found. Hence, we chose to look at other factors that could potentially be an indicator as to when terrorism would occur. One of these factors could be speech versus no speech, whether terrorists choose to communicate or stay silent before terrorist attacks. The graphs below explore the phenomena of speech versus no speech on the threshold value (the point at which the p-value[1] changes from significant to non-significant). As scores get higher (x-values), more intervals move from "high

[1] p-value for logrank calculated with R

score" category to the "no speech" category. The switching of intervals causes a decrease in the number of "high score" intervals to the right of the graph.

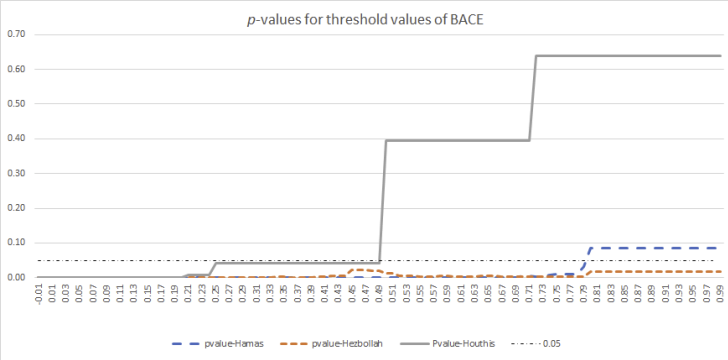

Figure 16. Log-rank *p*-values for threshold values of BACE.

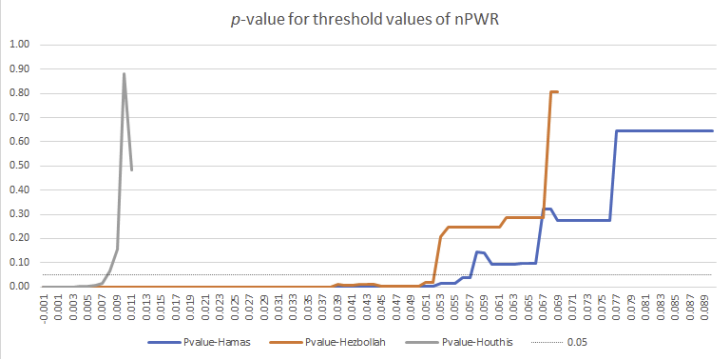

Figure 17. Log-rank *p*-values for threshold values of nPWR.

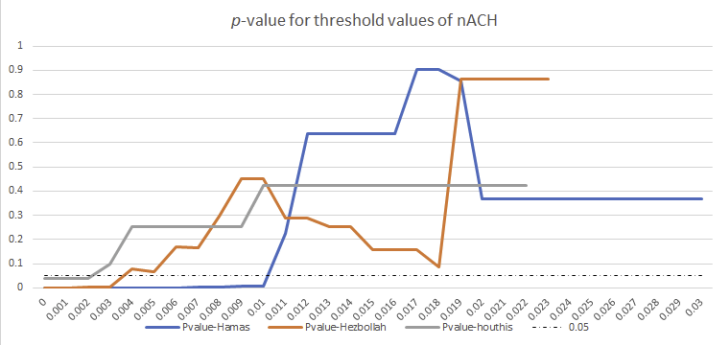

Figure 18. Log-rank *p*-values for threshold values of nACH.

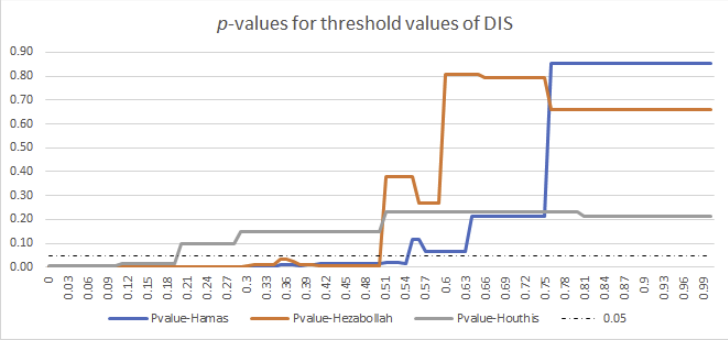

Figure 19. Log-rank *p*-values for threshold values of DIS.

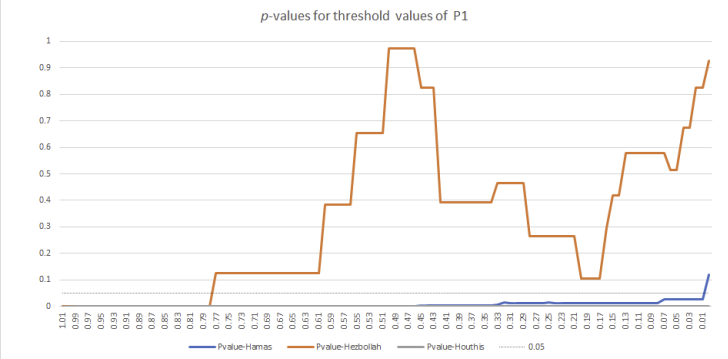

Figure 20. Log-rank *p*-values for threshold values of P1.

As displayed in Figures 16 through 20, many variables were significant until they reached a certain point (threshold value). As scores get higher (x-values), more intervals move from "high score" category to the "no speech" category. The switching of intervals causes a decrease in the number of "high score" intervals to the right of the graph. This decrease results in higher (non-significant) *p*-values towards the right end of the graph. This phenomenon shows the importance of speech versus no speech in our study. While most previous studies have explored the impact that psychological variables have on terrorism, few have looked at how the presence/absence of speech impacts the frequency of terror attacks.

Displayed in the BACE graph above (Figure 16), for both the Houthis and Hamas the difference between the "no speech" and "high score" categories loses significance as the threshold increases. The graphs for nPWR, nACH, and DIS for Hamas and Hezbollah follow a similar pattern as intervals are switched to the left or "no speech" category. In the nPWR graph, the Houthis have a bizarre cutoff in the graph, the reasoning behind this will be explored later

(we don't know what's happening!). In the nACH and DIS graphs, the Houthis follow a similar pattern to Hezbollah and Hamas. In the graph for P1, the direction of the x-axis is reversed to correspond to the hypothesis that lower P1 scores are associated with shorter intervals; intervals are shifted left to the "no speech" category as P1 decreases, leading to higher *p*-values for lower scores of P1.

Discussion

The results from our study have been surprising, as previous studies' results have not been supported—we do not find an association between differences in the psycholinguistic variables and length of interval between attacks. A partial explanation may be the impact of speech versus no speech, as most previous studies focused on the effect of differences in the speech variables on behavior but did not address the effect of the presence versus absence of speech. In those studies, focusing on the differences in the speech variables may have masked a stronger signal in the choice to communicate at all. In this study, we included all intervals between attacks regardless of the presence or absence of speech with the expectation that differences in speech variables would be associated with interval length. The initial results appeared to confirm this expectation but in the opposite direction.

The Houthis results were the most erratic for the three groups and this may be because the Houthis were engaged in active warfare during the time frame of our data collection, but Hamas and Hezbollah were not. Hence, there were thousands of attacks and for some periods attacks occurred every day for weeks. This made it exceptionally hard to find speech data, and there are many intervals for the Houthis that do not have speech data at all (631/660). Data for the Houthis from periods of intermittent conflict may give results more congruent with those for Hamas and Hezbollah.

The results for Hamas and Hezbollah are consistent across R and NCSS but still contradict every single hypothesis from previous studies. We found a consistent speech/no speech association with interval length for all five psycholinguistic variables which diminishes as intervals are reclassified as “no speech” as the threshold between the groups moves to the right. We have not ruled out a difference in the underlying psychological states of the terrorist groups during longer or shorter intervals, however that difference may manifest as a choice to communicate or not rather than manifesting in a change in the psycholinguistic variables. This may be a fundamental difference between terrorist or other non-state actors and national leaders who are often required to communicate with someone on a regular basis, especially in crisis or conflict situations.

Remaining issues

1. Most appropriate test for survival curve difference in this context.
2. Odd results for the Houthis nPWR results.
3. Discrepancies between NCSS and R for the Houthis.

Bibliography

Abdalsalam, M., Li, C., Dahou, A., & Noor, S. (2021). A Study of the Effects of Textual Features on Prediction of Terrorism Attacks in GTD Dataset. *Engineering Letters*, *29*(2). https://www.engineeringletters.com/issues_v29/issue_2/EL_29_2_11.pdf

Conway III, L. G., & Conway, K. R. (2011). The terrorist rhetorical style and its consequences for understanding terrorist violence. *Dynamics of Asymmetric Conflict*, *4*(2), 175-192

Conway, L. G., Gornick, L. J., Houck, S., Towgood, K. H., & Conway, K. R. (2011). The hidden implications of radical group rhetoric: Integrative complexity and terrorism. In *The Relationship between Rhetoric and Terrorist Violence* (pp. 71-81). Routledge

Hermann, M. G., & Sakiev, A. (2011). Leadership, terrorism, and the use of violence. *Dynamics of Asymmetric Conflict*, *4*(2), 126-134.

Hogenraad, R. L., & Garagozov, R. R. (2008). Words of swords in the Caucasus: About a leading indicator of conflicts. *Peace and Conflict: Journal of Peace Psychology*, *16*(1), 11

Huamaní, E. L., Alva, M. A., & Roman-Gonzalez, A. (2020). Machine learning techniques to visualize and predict terrorist attacks worldwide using the global terrorism database. *International Journal of Advanced Computer Science and Applications*, *11*(4).

Lazarevska, Elena, Jayne M. Sholl, and Michael D. Young. (2006) "Links Among Beliefs and Personality Traits: The Distinctive Language of Terrorists" in *Beliefs and Leadership in World Politics: Methods and Application of Operational Code Analysis.* Edited by Mark Schafer and Stephen G. Walker. Palgrave Macmillan.

Levine, Nick and Michael D Young (2014) “Leadership Trait Analysis and Threat Assessment with Profiler Plus”. Proceedings of ILC 2014 on 8th International Lisp Conference,Montreal, QC, Canada — August 14 - 17, 2014. Association for Computing Machinery.

Meng, X., Nie, L., & Song, J. (2019). Big data-based prediction of terrorist attacks. *Computers & Electrical Engineering*, *77*, 120-127.

Munk, T. B. (2017). 100,000 false positives for every real terrorist: Why anti-terror algorithms don't work. *First Monday*.

Pennebaker, J. W. (2011). Using computer analyses to identify language style and aggressive intent: The secret life of function words. In *The Relationship between Rhetoric and Terrorist Violence* (pp. 8-18). Routledge

Smith, A. G. (2004). From words to action: Exploring the relationship between a group's value references and its likelihood of engaging in terrorism. *Studies in Conflict & Terrorism*, *27*(5), 409-437.

Smith, A. G. (2008). The implicit motives of terrorist groups: How the needs for affiliation and power translate into death and destruction. *Political Psychology*, *29*(1), 55-75

Smith, A. G. (2011). The relationship between rhetoric and terrorist violence: Introduction. In *The Relationship between Rhetoric and Terrorist Violence* (pp. 1-7). Routledge

Walker, S. G. (2011). Anticipating attacks from the operational codes of terrorist groups. In *The Relationship between Rhetoric and Terrorist Violence* (pp. 51-59). Routledge

Winter, D. G. (2011). Scoring Motive Imagery in Documents from Four Middle East Opposition Groups. Dynamics of Asymmetric Conflict, 4, 144-154