The greatest number of victims in wars are civilians. Of the 50 million victims of World War II, only 20 million were military (Keegan, 1989). The Iraq War started in 2003 and produced over 100,000 deaths between 2004 and 2010, where 60,000 of these deaths were civilians (Dewar, 2010). This study aims to find who and what is responsible for these civilian deaths in the 2003 Iraq War. At the same time, it will illustrate how a new generation of free, open data analysis tools can empower any researcher to answer important questions about the state of the contemporary world. In most previous conflicts, civilian deaths were the product of random acts of violence rather than direct attacks. Therefore, we hypothesized that the Iraqi civilians very likely were to be killed by random acts of violence such as those created by improvised explosive devices (IEDs), originally set by Iraqi insurgents to kill American and allied troops. In order to test our hypothesis, we analyzed WikiLeaks’s Iraq War Logs, a dataset of 391,832 significant acts of war recorded by U.S. troops between November 6, 2004 and April 23, 2009. We used Python scripts, the R statistical analysis package, and Microsoft Excel to format, sort, and analyze the data. Our findings indicate that IED explosions contributed to 31% of civilian deaths, while direct fire contributed to 7% of civilian deaths. A comparison of how civilian deaths related to insurgent and allied intent shows that more civilians were killed by insurgents than by allied troops. Surprisingly, however, nonmilitary murder accounted for 49% of civilian deaths. A one-way analysis of variance (ANOVA) shows that the differences between these causes are strongly significant. Although the findings incompletely support the hypothesis, they reveal the complex nature of violence in Iraq and the multiple effects military intervention can have in a country.


Keywords

civilian death, IED, Iraq War, insurgents, murder, war logs, WikiLeaks
CIVILIAN DEATHS AND THE IRAQ WAR: Who Is Responsible?

Mihai Avram, Mathematics

INTRODUCTION

Modern technology has provided many deadly war weapons, from heat-seeking missiles, to improvised explosive devices, and to the most destructive of them all, nuclear weapons. Due to these deadly weapons, modern wars often have many casualties (Clark, 2002). Data about casualties indicate that it is civilians who are harmed the most by these weapons. The Iraq War is a vivid example of a major conflict in which civilian lives were lost. More specifically, between March 2003 and March 2013, approximately 120,000 civilians lost their lives (Dardagan, 2013). This study aims to identify who and what is responsible for this voluminous killing of civilians with the intent of extending the body of knowledge pertaining to military impacts in times of war. Furthermore, this study also illustrates how free, open data can be used to promote citizen journalism.

Literature Review

The Iraq War is one of the modern events that will remain in the annals of history. It changed the face of the Middle East and has reoriented U.S. foreign policy. The potential causes of the war were: oil, terrorism, the regulation of Iraq’s dictatorial regime, or a combination of these factors (Feeney, 2013). The United States intended to strategically control the Middle East for both military and economic reasons. Moreover, the United States has had historical political differences with Iran for the past few decades. Consequently, the United States planned to block Iran from the political and economic resources located in Iraq and to eliminate any possibility that the other major anti-U.S. actor in the region, Iraq, could use terrorists to get back at U.S. interests. Ensuring continued and free access to energy sources, such as oil, was also a reason for the invasion (Feeney, 2013). At the same time, the U.S. strategic goals needed to be put in perspective. The terrorist attacks of September 11, 2001 left the world shocked and the U.S. feeling vulnerable. Fearing a potential coalition of all possible U.S. political conflicts in the Middle East, with or without Al-Qaeda’s participation, the U.S. turned its policies from defensive to offensive. Furthermore, the U.S. government believed that Saddam Hussein, the president in Iraq at the time, was able to deploy weapons of mass destruction (WMDs). Consequently, the U.S. government requested Hussein to dispose of the WMDs. In the same context, Hussein had repeatedly failed to give a full account of how the weapons of mass destruction were disposed. Therefore, the U.S., as an advocate for democracy, found the amassing of WMDs by Hussein a threat to national and global security. Given that in the hands of a dictator such weapons could lead to unexpected outcomes, such as transfer to non-state actors like Al-Qaeda, the U.S. found reason (albeit contested by some as insufficient justification) to invade Iraq (Feeney, 2013). At the same time, Hussein’s policies were a potential threat, not only for the world, but they were a clear and present danger for his own people. The Iraqi nation, especially the Shia, was greatly oppressed by its government, and this created a latent religious and geographic conflict and regional instability (Hanley, 2005).

Figure 1 (Above). A U.S. soldier makes his way down the road as a canal burns in Tahwilla, Iraq where extremists used concealment provided by the intricate canal system to place IEDs under the cover of night. Courtesy of U.S. Department of the Army—photo taken by Spc. David Marshall.
In view of this synopsis, the war can be viewed from a broader perspective. After the September 11, 2001 attacks, the United States decided to invade Iraq in 2003 despite a lack of a United Nations mandate (Iraq War Logs, 2013). The basic strategy for the Iraq invasion was summarized as the “1% doctrine” (Suskind, 2006). By a cost-benefit analysis, the Bush and Cheney administration decided that even if the risk of weapons of mass destruction proliferation in the Middle East was 1%, it was worth declaring war against Iraq to prevent such a low probability. The United Kingdom, Australia, and many other nations sent troops to fight alongside the Americans, either due to shared strategic interests or desire to strengthen alliances (Iraq War Logs, 2013). Concurrently, Al-Qaeda also decided to join the war in Iraq, arming and financing militants like Abu Musab al Zarqawi (Cruickshank & Ali, 2007). The local insurgents included both Sunni and Shiite militants and other various groups who sought to fight for prominence and rights. The Sunni and Shiite militants revolted against the U.S. invasion hoping to rebuild a new government based primarily on their principles. In 2003, Hussein was captured and later hanged (Peterson, 2007). However, Hussein’s death did not put a stop to the killings. As the war raged between the allies, insurgents, and Al-Qaeda, there were about 8,803 military casualties and about 23,013 civilian casualties in 2006 and 2007 alone. Civilians were also slain as the Sunnis and Shiites fought for religious and political power (Hanley, 2005). Due to the large cost of the war and also because of Iraq’s growing political stability, the U.S. decided to remove troops from Iraq (Arnold, 2008; Biddle, Friedman, & Shapiro, 2012). Despite the positive effects of the U.S. offensive of 2007 and the subsequent withdrawal of the American troops, political and regional tensions in Iraq still exist today.

As mentioned, throughout these events civilian casualties were a constant outcome of the Iraq violence. There were select noteworthy events in which many civilians were killed: the Saddam Hussein processions, the Haditha massacre, the Samarra attack, and the Blackwater scandal (Iraq War Logs, 2013). In many of these instances, various distinct groups were responsible for civilian deaths. U.S./ allied troops, insurgents, Al-Qaeda, and Iraq troops were all responsible one way or another for civilian deaths. However, these historical accounts presented only a partial qualitative view of how civilians died in Iraq, and this quandary needed to be analyzed further by looking at the war from a quantitative point of view.

### Dataset

In October 2010, WikiLeaks, a nonprofit organization, the main objective of which is to bring important classified information to the public, released one of the largest classified military leaks in history (WikiLeaks, 2013). This leak is a dataset called the Iraq War Logs. This dataset contains 391,832 reports about events during the Iraq War as seen from the perspective of the U.S. military through Significant Activities (SIGACTS) reports (Bohannon, 2010). These events range from January 1, 2004 to December 31, 2009. The Iraq War Logs is the dataset on which this study is based.

### Previous Research

Previous research has investigated civilian deaths in Iraq; but identification of the exact causes of these deaths has been lacking. Amongst this research criterion, a study done between May 2006 and July 2006 concluded that most civilian deaths were due to violence—the most common type of violence being gunfire (Burnham, Lafta, Doocy, & Roberts, 2006). This study administered a mortality poll within 16 randomly selected governorates within Iraq, where each governorate contained approximately 2,000 households. The conclusions drawn from this study were valuable but it is limited in that it fails to incorporate who was at fault for the killings. Moreover, the events represented in this study covering 3 months are only a fraction of the full timeline of events covering the whole Iraq War.

Another similar study was conducted from 2003 to 2008, and it identified violent deaths of Iraqi civilians by perpetrator, weapon, time, and location (Hicks et al., 2011). The researchers used the Iraq Body Count (IBC) database, which is an organization responsible for “the worldwide update of civilian deaths in the Iraq war and occupation” (Dardagan, 2013). The researchers then used Stata, a data analysis statistical software, in order to analyze the database and find results for the
violent deaths of Iraq civilians. This study found that unknown perpetrators caused 74% of violent deaths of civilians, coalition or allied United States forces caused 12%, anti-coalition caused 11%, and military crossfire caused 2%. Moreover, the top two weapons that caused the most deaths were suicide bombings by unknown perpetrators and attacks by the coalition. A limitation with this study, however, is that this data does not have accurate reports on the exact weapons used for bombings or coalition attacks (Hicks et al., 2011).

Finally, a closely related study is the IBC database itself. Using an arsenal of about 90 news sources, including the Iraq War Logs dataset used in our study, IBC was able to attribute 71.8% of total civilian deaths to an unknown group, 12.1% to coalition forces, 14.6% to insurgents, and 1.5% to Iraqi forces. The Iraq War Logs dataset adds about 15,000 more civilian deaths to the IBC database and a keener, more detailed interpretation of events in Iraq. Although the Iraq War Logs dataset is merely a subset of IBC, with careful analysis of the War Logs, a more exact turn of events, group, and weapon can be established to be responsible for the civilian deaths in Iraq.

**Research Questions**

In utilizing the Iraq War Logs dataset, we are now in the position to ask better questions and provide better answers about the causes of the civilian deaths in Iraq. Although limited, as the dataset only records data controlled by the U.S. Army, it is one of the most complete records to date of the civilian deaths in Iraq (WikiLeaks, 2013). The Iraq War topic is a relatively new topic that has not been studied too rigorously yet. Consequently, some questions still remain unanswered regarding civilian deaths in Iraq. This study seeks to answer the following questions: Who was responsible for killing helpless civilians? And
what weapon or method of attack was responsible for this outcome, and under what circumstances?

**Variables**

The dataset (Figure 3) acquired from the WikiLeaks website consists of 391,832 rows, each representing a “significant act of war” reported as a field brief by a U.S. military unit. For each act, the following information was recorded and stored in separate variables. The first two variables, type of incident and category, are the most important for our study. Type of incident distinguishes between criminal acts, activities initiated by the U.S. troops, enemy actions, or explosive hazards (mostly IEDs set by the enemy). Category of incident provides more in-depth information about each type of incident. For example, enemy actions can be categorized as attacks, ambushes, or indirect fire, while friendly actions can be categorized as attack, patrol, offensive, defensive, and so on. Each event has a death toll associated with it. Moreover, for each act, the number of killed and wounded is broken down into categories: civilian killed in action, enemy (insurgents) killed in action, friendly troops (U.S. and allies) killed in action, and host nation (Iraqi troops allied with the U.S.) killed in action. Thus, it is possible to determine how many civilians were killed in criminal acts or due to U.S. or enemy attacks. The time and the precise location in latitude and longitude for each event also have been recorded.

**METHODS AND RESULTS**

We hypothesized that the civilian deaths in the Iraq War would be the largest of all group deaths. In order to examine this hypothesis, we needed to quantify the percentages of groups killed in action during the
war. First, we used Python, a computer programming language, to format the dataset for analysis. We then extracted the data about individuals killed in action broken down into military and civilian groups, and using the statistical computing software, R, we were able to compare these datasets amongst each other to find percentages of each group killed.

We found that of all personnel killed during the war, 3% were U.S. and allied troops, 14% were Iraqi troops, 22% were insurgents, and 61% were civilians (Figure 4). This validated our hypothesis that civilians were the largest group killed in the war. In the next step of our investigation, we sought to identify the groups and events responsible for the civilian deaths. Consequently, we needed to isolate civilians killed in action and examine the specific events that contributed to these civilian deaths. During preliminary analysis using Microsoft Excel, we observed that the top five events contributing to 93% of the civilian deaths were, in order: murder, IED explosion, direct fire, indirect fire, and attack. In further investigation, we used R to analyze specific events and were able to attribute exact death counts of civilians for each event.

We established that 32,564 civilians were murdered, 20,326 civilians were killed by IED explosions, 4,767 civilians were killed by direct fire, 2,087 were killed by indirect fire, and 1,628 were killed by attacks (Figure 5). Although the Iraq War Logs do not specifically define these events, we closely examined qualitative records of the individual logs that contained specific information for each event (murder, IED explosion, direct fire, indirect fire, and attack). Results found that 49% of civilians were murdered by gunfire, explosions, drowning, stabbings, and assaults. IEDs accounted for 31% of civilian deaths; 7% of civilians were killed by direct fire from firearms; another 3% of civilians were killed by indirect fire from firearms; and finally, 2.5% of civilians were killed by attacks or assaults. After identifying the events that were responsible for the majority of civilians dying, we sought to identify the factions responsible for these events. To approach this problem, we used R to tally the number of civilian deaths by broad types of action.

The first and most interesting conclusion was that 50% (32,983) of the 66,081 deaths recorded by the SIGACTS U.S. military reports were the product of criminal activity (Figure 6). While the definition of criminal activity is unclear, generally it describes an act of violence that does not directly involve an easily identifiable political or military cause or direct involvement of either U.S. or insurgent personnel. At the same time, criminal acts need to be understood in context. It is possible that many of them would be acts of revenge or violence motivated in one way or another by political or strategic reasons that are not immediately visible. More important for our investigation, however, is whether the tallying of civilian deaths is attributable, according to U.S. military sources, to U.S. versus enemy combatant activities. In this respect, there is a one to five ratio in favor of the enemy combatants. For each civilian killed by U.S., Iraqi, or other troops, there were five civilians killed by insurgents. Furthermore, if we focus our attention on the 31,042 deaths that can be directly attributable to the U.S. or friendly troops versus enemy combatants, it appears that 95% were due to enemy combatant activities, according to the U.S. field reports (Figure 7). This is representative of about half of all civilian deaths during the Iraq War.

However, this is a very rough estimate that takes U.S. military reports at face value. We decided to take an alternative route for determining possible connections between civilian deaths and war activities. Our general approach was to determine if, for events that were categorized neutrally and possibly deceivingly as

<table>
<thead>
<tr>
<th>Death Variables Correlation–IED</th>
<th>Correlation</th>
<th>P Value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civilians Killed in Action vs. Iraqi Troops Killed in Action</td>
<td>0.05</td>
<td>&lt;0.001</td>
<td>Very Significant</td>
</tr>
<tr>
<td>Civilians Killed in Action vs. Insurgents Killed in Action</td>
<td>0.07</td>
<td>&lt;0.001</td>
<td>Very Significant</td>
</tr>
<tr>
<td>Civilians Killed in Action vs. United States/Allies Killed in Action</td>
<td>0.00</td>
<td>0.57</td>
<td>Not Significant</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Death Variables Correlation–Murder</th>
<th>Correlation</th>
<th>P Value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civilians Killed in Action vs. Iraqi Troops Killed in Action</td>
<td>-0.16</td>
<td>&lt;0.001</td>
<td>Very Significant</td>
</tr>
<tr>
<td>Civilians Killed in Action vs. Insurgents Killed in Action</td>
<td>-0.02</td>
<td>&lt;0.001</td>
<td>Very Significant</td>
</tr>
<tr>
<td>Civilians Killed in Action vs. United States/Allies Killed in Action</td>
<td>-0.02</td>
<td>0.01</td>
<td>Significant</td>
</tr>
</tbody>
</table>

Table 1. Correlations between civilians killed in action and groups: Iraqi troops killed in action, insurgents killed in action, and US/allied troops killed in action. These correlations are done with respect to IED explosions and murder.
“murder” or “IED,” the number of civilian deaths was related to U.S. and allied troops, Iraqi troops, or insurgent deaths in some way. The logic is simple: if the number of civilian dead increases, while that of U.S. and friendly troops decreases, then there could be some evidence that the “killing game” is one in which one party “wins” and the other “loses.” Put another way, if the correlation between civilians killed in action and any other group killed in action is low or negative, then as more civilians die, the members of the other group are far less likely to die. Correlation quantifies how two variables are related or dependent upon each other. When the correlations are negative and significant there is also the possibility, which needs to be fully demonstrated through other means in further research, that the paired group (U.S./friendly troops or enemy combatants) is responsible for killing the civilians. We need to add that this analysis is conducted on events that are overly broad in nature and hard to pinpoint as to their true cause, namely, murder and IED explosion. Their poor definition can also be connected to reasons of self-preservation for the reporting agents. Note that this approach was based on intuition and needs to be further researched for validity.

Association between group deaths was assessed via Pearson’s r correlation. The pairs of variables that were correlated were: civilians killed in action versus enemy combatants (insurgents) killed in action; civilians killed in action versus Iraqi troops killed in action; and civilians killed in action versus U.S./allied troops killed in action. Correlations control for cause of death, namely, murder and IED explosion. While the results do not provide direct evidence for direct causality, since correlation does not imply causation, it does shed light on possible trends in the data that deserve further investigation.

These correlations conducted in R (Table 1) are very low, and in conducting an analysis of variance (ANOVA) between the groups, we also found a significance level for each correlation. We used the threshold of 0.05, which means that if any two compared variables are dependent on each other more than 95% of the time, the correlation is significant. The most obvious and significant result is that there is a negative correlation between civilians and Iraqi troops killed in action when the context of the killing was murder. In other words, murder events tend to have more civilian deaths when there are fewer Iraqi soldier deaths.

While correlation is no causation, the negative balance sheet for civilians in the context that Iraqi forces are not affected suggests a trend that needs to be further explored (Figure 8). Although some of the other correlations are statistically significant, they are 5 to 10 times smaller in magnitude compared to Figure 8 and are too weak to suggest any meaningful relationship.

**Study Limitations**

Although this study has enriched the body of knowledge about civilian deaths in the Iraq War, there are some limitations. The first limitation is associated with the correlation tests between civilians killed in action and other groups, controlled for murder and IEDs. The correlation tests are not an accurate measure to base any

![Figure 8](image8.png)

*Figure 8.* Correlation graph between civilians killed in action and Iraqi troops killed in action controlling, for murder.

![Figure 9](image9.png)

*Figure 9.* U.S. soldiers provide crowd control over Iraqi civilians during a medical civic assistance program at the Basra Operations Center in Basra, Iraq. Courtesy of U.S. Department of Defense.
group’s involvement in civilian killings just by comparing overall deaths for each group. Another pitfall is that during times of war, it is hard to distinguish the essential intent of one group. An individual could have used explosives in two large trucks in a populated area with the intent of killing allied troops, but ended up killing Iraqi troops and civilians. Therefore, given the uncertainty of the dataset, we are not able to draw an exact conclusion about who was at fault. Also, the Iraq War Logs dataset does not include information about events that occurred in 2003, when the war commenced, or in 2010, when the war terminated, and hence does not include all the civilian deaths that occurred in Iraq. Finally, the Iraq War Logs were recorded by U.S. troops who witnessed the events, and as with any qualitative reports, there may be inaccuracies.

**CONCLUSIONS**

In summary, previous research studies have investigated Iraq War casualties; however, the investigation of specific causes for civilian casualties is a significant research gap. This study aimed to identify events during the Iraq War that contributed to the deceased civilians. Furthermore, this study also sought to identify the particular weapons, events, and militant groups responsible for these civilian deaths. In conclusion, we found the following to be the partial causes of civilian deaths: gunfire, explosions, drowning, stabings, and assaults (49%); IED explosions (31%); direct fire from firearms (7%); indirect fire from firearms (3%); and attacks or assaults (2.5%). Additionally, the types of events responsible for 50% of civilian deaths were criminal events motivated by political or strategic reasons. Moreover, according to U.S. troops, for every civilian killed by U.S., Iraqi, or other troops, there were five civilians killed by insurgents. Additionally, of the 31,042 deaths that can be directly attributable to the U.S. or friendly troops versus enemy combatants, 95% of the deaths were due to enemy combatant activities. Therefore, enemy combatants are responsible for a far larger number of civilian deaths than U.S. troops. Future research studies using the Iraq War Logs dataset should examine each event out of the 391,832 different logs in the dataset and be able to distinguish what factions were responsible for these events that resulted in the killing of civilians. Details regarding clashes between Iraqi troops and civilians also should be investigated further.

**REFERENCES**


---

**Figure 10.** Aftermath of a vehicle-born IED in Iraq. Courtesy of U.S. Department of Defense—photo taken by a member of the 187th Infantry Regiment.