How Combining Terrorism, Muslim, and Refugee Topics Drives Emotional Tone in Online News: A Six-Country Cross-Cultural Sentiment Analysis

CHUNG-HONG CHAN
HARTMUT WESSLER
University of Mannheim, Germany

EIKE MARK RINKE
University of Leeds, UK

KASPER WELBERS
WOUTER VAN ATTEVELDT
Vrije Universiteit Amsterdam, the Netherlands

SCOTT ALTHAUS
University of Illinois at Urbana-Champaign, USA

This study looks into how the combination of Islam, refugees, and terrorism topics leads to text-internal changes in the emotional tone of news articles and how these vary across countries and media outlets. Using a multilingual human-validated sentiment analysis, we compare fear and pity in more than 560,000 articles from the most important online news sources in six countries (U.S., Australia, Germany, Switzerland, Turkey, and Lebanon). We observe that fear and pity work antagonistically—that is, the

Chung-Hong Chan: chung-hong.chan@mzes.uni-mannheim.de
Hartmut Wessler: wessler@uni-mannheim.de
Eike Mark Rinke: E.M.Rinke@leeds.ac.uk
Kasper Welbers: kasperwelbers@gmail.com
Wouter van Atteveldt: wouter@vanatteveldt.com
Scott Althaus: salthaus@illinois.edu
Date submitted: 2019-08-30

1 This project was funded by a research grant from the (1) German Research Foundation (Deutsche Forschungsgemeinschaft), (2) the Netherlands Organisation for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek), and (3) the National Endowment for the Humanities, through the Trans-Atlantic Platform’s Digging Into Data Challenge funding program. The authors would like to thank the following coders and translators (in alphabetical order): Mohamad Alsaleh, Gamze Bozkurt, Rainer Freudenthaler, Patrik Haffner, Mohammad Joudeh, Büşra Karaca, Katharina Ludwig, Michael Martin, Matthew J. Powers, Sarah Shammaa, and Helin Yilmaz.

Copyright © 2020 (Chung-Hong Chan, Hartmut Wessler, Eike Mark Rinke, Kasper Welbers, Wouter van Atteveldt, and Scott Althaus). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at http://ijoc.org.
more articles in a particular topical category contain fear, the less pity they will feature. The coverage of refugees without mentioning terrorists and Muslims/Islam featured the lowest fear and highest pity levels of all topical categories studied here. However, when refugees were covered in combination with terrorism and/or Islam, fear increased and pity decreased in Christian-majority countries, whereas no such pattern appeared in Muslim-majority countries (Lebanon, Turkey). Variations in emotions are generally driven more by country-level differences than by the political alignment of individual outlets.

**Keywords:** terrorism, refugee, Muslim, sentiment analysis, multicultural analysis

Human migration is extensively covered in the media, and how the media report on the topic of refugees is of great scholarly interest. In 2015–2016, refugees, who were mostly Syrians displaced from their war-torn country, migrated to neighboring countries such as Turkey and Lebanon as well as to Europe. This massive human migration led to multiple cultural and political consequences. Media in the receiving countries played an important role in framing the influx of refugees and often took to framing refugees negatively, as a threat in various ways (Eberl et al., 2018). Several studies have compared how the topic of refugees was reported in terms of tone and framing in European countries in 2015–2016 (e.g., Berry, Garcia-Blanco, & Moore, 2016; Chouliaraki & Zaborowski, 2017; Heidenreich, Lind, Eberl, & Boomgaarden, 2019). These studies rarely consider how the coverage of refugees in European countries compares to that from media outlets in non-European countries, especially countries that were (1) much more affected by the refugee movements of 2015–2016 and (2) more culturally close to the involved refugee cohorts. We argue that such cross-cultural comparison is essential if we are to understand how culture plays out in shaping the coverage of terrorism, Muslims, and refugees. Therefore, this study examines the emotional tone that follows from combining the topics of refugees, terrorism, and Muslims/Islam in reportage from 37 online news outlets located in six Western and non-Western countries. We thus add a cross-cultural, comparative perspective to the study of migration in the media and thus put existing theories about refugee coverage to a test outside the Western world. In the following sections, we unpack the main components of our analysis: (1) combining the topics of terrorism, Muslims, and refugees (TMR); (2) the emotions in TMR coverage; and (3) the mechanisms that help explain differences in emotional tone.

**Combination of Topics**

In 2015–2016, one can observe three streams of events: the influx of refugees, the integration of Muslims into the Christian-majority European countries, and a wave of worldwide Islamist terrorist attacks. This concurrency provides journalists the opportunity to combine the three topics because some aspects of the three events are related. For example, many refugees are Muslims, and radical Islamist terrorists have attacked European countries in the name of Islam. Instead of studying refugees as a singular discourse (such as the studies included in the review by Eberl et al., 2018), the combinations of TMR topics in news coverage need to be investigated to understand how combination alters the emotional tone that would be applied to each topic alone.
When we study the three topics simultaneously, it enables us also to study how suspect communities are constructed. The term *suspect community* was first coined to describe minority communities that are associated with the perpetrators of terrorist attacks based on public characteristics such as religion (Pantazis & Pemberton, 2009). The underlying logic is to discursively combine two communities with similar or overlapping characteristics. For example, many refugees are Muslims, and radical Islamist terrorists are also Muslims. Some political actors might combine the two, and thus the refugee community is cast as a suspect community, an interpretation that might then infuse media coverage. Associating Muslims and refugees with terrorism is particularly problematic because experimental evidence shows that this can trigger more intense fear reactions in readers and leave some readers with a more hostile attitude toward the nonterrorist (i.e., Muslim) groups (von Sikorski, Schmuck, Matthes, & Binder, 2017).

In practice, Western media sometimes do not differentiate the three topics. This practice of combining TMR topics likely contributes to a more negative public perception of Muslims and refugees. Muslims were usually negatively reported in the U.S. media (e.g., Bleich, Stonebraker, Nisar, & Abdelhamid, 2015; Powell, 2011). A meta-analysis showed that Western media usually apply negative frames to Muslims (Ahmed & Matthes, 2016). In the U.S. media’s reporting on the 9/11 attacks, the event was mostly framed as a Huntingtonian clash of civilizations between Islam and the West (Abrahamian, 2003; Kellner, 2004). This frame emphasizes fundamental differences among cultures and is likely to widen the perceived distance between Muslims and non-Muslims. A critical discourse analysis showed that U.S. media frequently associate Muslims with violence, religious radicalism, and Islamist terrorists (Samaie & Malmir, 2017).

In summary, public discourses in Western media related to terrorism, Muslims, and refugees from Muslim-majority countries are sometimes intertwined. However, previous studies usually investigate the three topics as separate topics. Research on the combination of the three topics is sparse. Most studies are from the era before 2015 and do not deal with refugees. We argue, therefore, that we must study the three topics as intersecting if we hope to understand their discursive dynamics and consequences. This study sets out to make this point by tracing the systematic variation of emotional tone when these three topics are combined in media coverage.

**Emotional Tone Variations**

The tone of news coverage is defined as the general positive and negative valence of a news item (e.g., Bleich et al., 2015). The study of emotional tone (or simply emotions in news) goes beyond this general dichotomy by investigating to what degree distinct emotions are expressed in the news (Cho et al., 2003). Some emotions are positively valenced, such as joy; some carry negative valence, such as sadness, anger, or fear; and some are mixed, such as pity, which combines the negative valence of

---

2 For example, UK and U.S. mainstream media reported the 2005 failed London bomb plot as "extremist Muslim plot" (http://edition.cnn.com/2007/WORLD/europe/01/17/warwithin.overview/). In another example, Fox News reported the refugee situation in Germany with a paragraph about the rising Muslim population in Europe (https://www.foxnews.com/world/germany-offers-money-for-migrants-to-go-back-home).
somebody’s suffering with the positive valence of sympathy for the sufferer. The study of emotional tone in news items is a subfield of what is usually called sentiment analysis (Cho et al., 2003). News items can contain traces of several emotions to varying degrees at the same time. In our study, we examine how the prevalence of each emotion is changed by the topical aspect(s) covered in the item.

The main challenge in measuring emotional tone lies in determining how explicit the emotion is expected to be expressed. For example, pity could be expressed by a government spokesperson quoted with the words, “We express our sincere condolences to the victims of the attack.” This is rather explicit. On the other hand, the item could also quote somebody as saying, “The human tragedy in Syria has forced so many of our countrymen to leave their home and their loved ones.” Even though the quote does not express a feeling of pity, the sentence’s content describes something that the speaker feels sorry about. This, therefore, is an implicit expression of pity. Both explicit and implicit expressions of pity (and, by extension, of any emotion) can be measured by capturing word choice. In the explicit quote, the words would be, for example, *condolences* and *victims*; in the implicit case, they would be *tragedy* and *forced*.

Past work on the emotional tone of terrorism and Muslim coverage has focused on negative emotions such as fear and anger, not least because mass-mediated fear is a key objective of terrorist attacks (Cooper, 2001). Emotional tone in TMR coverage has also been shown to generate effects on media users’ psychological response (Uribe & Gunter, 2007), attitudes, and behaviors (Nabi, 2003). Using automated text analysis, Cho et al. (2003) and Boomgaarden and Vliegenthart (2009) observed differences in emotions of TMR coverage and readers’ responses, such as fear and anti-immigration attitudes.

In the Muslim context, the emotional tone of TMR reporting is important in yet another way. Radicalization is both the cause (McCauley & Moskalenko, 2008) and effect (P. R. Neumann, 2013) of terrorist attacks. In a laboratory experiment (K. Neumann, Arendt, & Baugut, 2018), negatively framed news about Muslims elicited anger and intensified hostile media perceptions among Muslim subjects living in Western countries. These factors also contribute to Islamist radicalization.

In contrast to the dominant discourse on Muslims, the coverage of refugees is characterized more by sadness and pity as major emotions (e.g., Chari, 2010; Höijer, 2004). But nonnegative emotions such as pity have only been studied sparingly in the TMR context, even though rescue and help for victims or refugees offer at least some room for expressions of pity in the media.

To sum up, emotions play a crucial, and potentially very consequential, role in media coverage of terror, migration, and outgroups more generally. Therefore, this study goes beyond the usual binary classification of negative versus positive news coverage by focusing on two particularly important discrete emotions in the TMR context: namely, fear and pity.³

³ We originally aimed at also studying anger, sadness, and joy, and we report the results for these three emotions in the online appendix (https://doi.org/10.17605/OSF.IO/A4DQP). Theoretically, anger, sadness, and joy are much less important in the context of TMR coverage than fear and pity. Pragmatically, there is large variation in the accuracy of our customized dictionaries across the five emotions, and the
Fear was found to be the dominant emotion expressed in terrorism coverage in the U.S. and Russia (Oates, 2006). Some politicians in Western countries have long used fear-laden news coverage of terrorism in political campaigns (e.g., Oates, 2006). When mixed with other topics such as Muslims, one should expect that coverage of terrorism will be laden with fear even more and create negative perceptions (Powell, 2011).

Pity, on the other hand, is the major sentiment expressed in refugee coverage (Chari, 2010; Höijer, 2004). Covering refugees is a form of mediated representation of suffering, either remotely or domestically. Applying Chouliaraki’s (2006) model of media discourse of suffering, there are three types of news about suffering and the media discourse of pity: (1) adventuristic news, which reports those who suffered as outsiders and blocks emotional engagement with them; (2) emergency news, which also reports those who suffered as outsiders, but provides readers with a possible course of action; and (3) ecstatic news, which reports those who suffered as people similar to us. Chouliaraki suggests that news in the last category is more likely to express pity. However, refugee coverage with mixed topics (e.g., refugees mixed with terrorism) may shift the ecstatic coverage of suffering into emergency or even adventuristic modes of reporting.

Mechanisms Influencing Emotional Tone of TMR Coverage

Although some previous studies have identified the varying emotions expressed in TMR reporting, we know little about the origins of those differences. Next, we summarize three possible mechanisms, on both the level of media outlets and the level of countries and cultures. Because these mechanisms have not been explicitly related to levels of fear and pity in media coverage, we derive research questions rather than hypotheses for each of the mechanisms.

**Deviance Mechanism**

Several studies show that the mechanism for the media to report on a community negatively is based on the relative distance of the targeted community to relevant mainstream values of the majority culture (e.g., Benson & Saguy, 2005; Nickels, Thomas, Hickman, & Silvestri, 2012). In non-Muslim-majority countries, we might thus observe more negative and less positive/mixed emotions associated with Muslims as well as combined topics involving Muslims. Conversely, media in Muslim-majority countries might exhibit more positive/mixed and less negative emotions when talking about Muslims. Applying these expectations to pity and fear, we arrive at the following research questions:

RQ1.1: Are Muslims and combined topics involving Muslims reported with a higher level of fear in non-Muslim-majority countries than in Muslim-majority countries?

RQ1.2: Are Muslims and combined topics involving Muslims reported with a lower level of pity in non-Muslim-majority countries than in Muslim-majority countries?

Dictionaries for fear and pity perform slightly better than the other ones, with two exceptions, which are discussed later.
Ideological Position Mechanism

There also is variation within Western media regarding their TMR reporting. For example, Fox News stood out as exceptional in the coverage of the “war on terrorism.” This finding was attributed to Fox’s ideological position relative to other U.S. media (e.g., Aday, 2010; Aday, Livingston, & Hebert, 2005). In the 2015–2016 refugee debate, the coverage of refugees in Europe was also found to vary by the political position of media outlets (Berry et al., 2016). For example, in Germany, the center-right newspaper Die Welt and the populist-right tabloid Bild framed refugees in more negative terms than the center-left Süddeutsche Zeitung. In our cross-cultural comparison, we might thus observe more negative emotional tone in the TMR reporting from right-wing media than from others. Hence, our research questions:

RQ2.1: Are Muslims and combined topics involving Muslims reported with more fear in right-wing news media than left-wing news media?

RQ2.2: Are Muslims and combined topics involving Muslims reported with less pity in right-wing news media than left-wing news media?

Regional Difference and Geopolitical Conflict Mechanism

Another relevant stream of studies looks at framing and argues that the reasons for media outlets to report TMR differently are based on regional differences in journalistic style. Papacharissi and de Fatima Oliveira (2008), for example, attribute the difference between UK and U.S. coverage of 9/11 to the observation that the U.S. press is more locally focused than the UK press.

Other studies attribute differences in TMR reporting to patterns of geopolitical conflict. For example, Gerhards and Schäfer (2013) contrast the coverage of four terrorist events in CNN, BBC, Al Jazeera, and German public broadcaster ARD. The European outlets (BBC and ARD) framed the attacks as crimes against humanity, whereas CNN and Al Jazeera framed them as manifestations of a geopolitical power struggle. The coverage of the latter two outlets differed in characteristic ways because they were situated at the two ends of that geopolitical conflict. Subsequent studies also suggest that geopolitical conflicts may influence how newsmakers report about terrorism and refugees (e.g., Lawlor & Tolley, 2017).

Apart from these theoretical expectations, we argue that the sociopolitical circumstances in the countries at the time of investigation can also shape emotions in the coverage of TMR topics. For example, journalists in countries with higher influx of refugees in 2015–2016 (e.g., Germany, Lebanon, and Turkey) might have different impressions of the refugees from Muslim-majority Syria and Afghanistan. Compared with countries more remote to the issue, these three countries experienced potential challenges in cultural integration and counterterrorism. Germany, for example, adopted an initial welcoming culture (“Willkommenskultur”) toward refugees. The German media might thus have shown, at least initially, a greater expression of pity toward refugees. At that moment also, the anti-immigration party Alternative für Deutschland was not as popular as it became later. Instead, anti-immigration parties in other Western
countries such as Schweizerische Volkspartei (Switzerland) and Australia First Party (Australia) were established political forces at the time and have a longer history of media mobilization (McGann & Kitschelt, 2005). In these countries, therefore, media reporting might have exhibited a more hostile attitude toward refugees on average. For instance, previous research on Australia shows that the media there have echoed the government’s negative tenor toward refugees (Klocker & Dunn, 2003).

How the media in Muslim-majority countries such as Lebanon (54% Muslims) and Turkey (83%) cover refugees from neighboring Syria and related issues is severely underresearched. It is generally believed that the media in these countries would report differently and might exhibit a higher level of pity and a lower level of fear than media in Western countries. This might also be reflected in the fact that displaced Syrians in Lebanon are not called refugees by the Lebanese media, but simply “Syrians.” Such circumstantial evidence is, of course, not sufficient to exactly predict whether the news media in any one country might exhibit a greater or lesser extent of fearmongering or compassion toward refugees. However, we anticipate that there should be some difference in the emotional tone of TMR coverage on the level of individual countries that reflect these specific circumstances (for the same conclusion, see Gerhards and Schäfer, 2013). Thus, we derive an exploratory RQ to study the individual country differences in coverage of TMR.

**RQ3:** To what extent does the coverage of terrorism, Muslims, and refugees in different individual countries exhibit differences in fear and pity?

**Research Gap**

As mentioned, almost none of the existing studies has examined media outside the Western world. Exceptions are few and far between (e.g., Aday et al., 2005; Gerhards & Schäfer, 2013; Wessler & Adolphsen, 2008). The few studies of media content from the Muslim world have often been restricted to the Al Jazeera television network. Even more often, they have eschewed comparisons of such content to media content found in other, non-Muslim cultures (e.g., Qadir & Alasuutari, 2013). But with a strictly intracultural comparative design, it is impossible to verify the aforementioned mechanisms comprehensively. For example, the ways in which mainstream cultural values operate in the deviance mechanism cannot be verified in Western-only settings. Therefore, this study includes news media outlets from countries with both predominantly Christian and predominantly Muslim traditions to examine how cultural in-group and out-group dynamics affect the emotional tone. We also deploy a research design that allows us for the first time to investigate all the mentioned mechanisms in a single study.

**Methodological Design**

This study is exploratory in nature. It strives to answer research questions and to accumulate unanticipated insights to generate further hypotheses for future studies. To this end, it undertakes a systematic description of emotional tone in three topics and their combinations and investigates sources of variation in emotional tone.
Data

The Mannheim International News Discourse Data Set (Rinke, Löb, Dobbrick, & Wessler, 2019) is a corpus of articles and posts from major news outlets and political blogs in six countries (U.S., Australia, Germany, Switzerland, Turkey, and Lebanon) during the period from August 1, 2015, to July 31, 2016 (for details, see Rinke, Zirn, Löb, & Wessler, 2017). For this study, we use online news only ($N = 567,925$). This data set is suitable for this study for three reasons: (1) The time period includes multiple terrorist attacks (e.g., Paris attacks in November 2015) and incidents inducing public discussions of multiculturalism (e.g., the 2015–2016 turn-of-the-year Kölner Silvesternacht incident during which numerous cases of sexual assault were committed in Cologne, Germany, mostly by migrants from Northern Africa); (2) this data set was collected from both Muslim- and Christian-majority countries; and (3) the choice of news outlets was based on expert surveys of media researchers in each of the six countries to represent the most important news outlets across the political spectrum.

Independent Variables

The fundamental independent variable is the topic of an article—that is, the fact that an article talks about terrorism (T), Muslims (M) or refugees (R), or any combination of these (T, M, R, TM, MR, TR, and TMR). The keywords used to identify articles (Table 1) were determined by native speakers involved in the project. For each language, a native speaker manually coded 250 articles in the eight categories (the mentioned seven combinations and articles without any TMR keywords) for their relevance to the topics of terrorism, Muslim(s), or refugees. The keywords have reasonable precision and recall with the exception that precision of our M keywords is less than desirable in the Arabic and German cases. This is because our native speaker coders did not code the bare mentions of organization names, such Islamic State, as related to Muslims; these mentions occurred more often in Arabic and German news articles. Our keyword-based approach is thus not perfect, but should be deemed adequate for an exploratory, hypothesis-generating study because almost all articles judged by humans to be related to Muslims were also captured by our keywords.

For each included article, we randomly selected another article as a control from the same date and the same media outlet without any TMR keywords. These controls were used to adjust for the base rate of emotions in coverage as determined by our sentiment analysis.

A second set of independent variables relate to the explanatory mechanisms specified earlier in the article. These variables were country, political alignment of the media outlet, and dominant religion. We summarize the categorizations of online media outlets by country and political alignment in Table 2. The political alignments of media outlets were coded into three categories by the third author and confirmed by the first and second authors.
Table 1. Precision and Recall of Search Keywords by Language.

<table>
<thead>
<tr>
<th>Search keywords</th>
<th>Precision %</th>
<th>Recall %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T (terror[tm]</td>
<td>terror attack)</td>
<td>87.0</td>
</tr>
<tr>
<td>M (muslim</td>
<td>islam)</td>
<td>68.0</td>
</tr>
<tr>
<td>R (refugee)</td>
<td>92.0</td>
<td>92.0</td>
</tr>
<tr>
<td><strong>German</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T (terror*)</td>
<td>88.0</td>
<td>82.2</td>
</tr>
<tr>
<td>M (muslim</td>
<td>moslem</td>
<td>islam</td>
</tr>
<tr>
<td>R (fl[uü]cht*)</td>
<td>65.0</td>
<td>92.9</td>
</tr>
<tr>
<td><strong>Turkish</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T ([tT]erörizm</td>
<td>[Tt]erörist</td>
<td>[tT]erör saldırlar</td>
</tr>
<tr>
<td>M ([Ill][a][Mm]üslüman)</td>
<td>81.0</td>
<td>97.5</td>
</tr>
<tr>
<td>R ([Mm]ülteci</td>
<td>[Mm]uhacir</td>
<td>[Ss]ığınan kimse)</td>
</tr>
<tr>
<td><strong>Lebanese Arabic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T (بیانیه: ایپه</td>
<td>مکت</td>
<td>باهرانلا</td>
</tr>
<tr>
<td>M (ساس</td>
<td>ای</td>
<td>باء</td>
</tr>
<tr>
<td>R (ذئ</td>
<td>ل</td>
<td>ر</td>
</tr>
</tbody>
</table>

Table 2. Categorization of Online Media Outlets.

<table>
<thead>
<tr>
<th>Country</th>
<th>Left/Center-left</th>
<th>Center</th>
<th>Right/Center-right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>Bianet, Cumhuriyet, OdaTV, Radikal</td>
<td>Hürriyet, Milliyet</td>
<td>Zaman</td>
</tr>
<tr>
<td>Lebanon</td>
<td>Al Akhbar, Al Jadeed, Lebanon Debate</td>
<td>An Nahar, Elnashra</td>
<td>Lebanon Files, Tayyar</td>
</tr>
<tr>
<td>Germany</td>
<td>Süddeutsche Zeitung, Zeit</td>
<td>Spiegel, Tagesschau</td>
<td>Bild, Frankfurter Allgemeine</td>
</tr>
<tr>
<td>Switzerland</td>
<td>–</td>
<td>20 Min, Neue Zürcher Zeitung, Schweizer Radio und Fernsehen, Tagesanzeiger</td>
<td>Blick</td>
</tr>
</tbody>
</table>
Dependent Variables

The dependent variables are sentiment levels of fear and pity in the online news articles. To quantify the emotions, we developed our own sentiment dictionaries, validated and adapted to TMR coverage (see Figure 1; refer to the online appendix for details).²

Figure 1. Dictionary development process for TMR articles in one language. Note: NRC = NRC Word-Emotion Association Lexicon; MFD = Moral Foundation Dictionary

² https://doi.org/10.17605/OSF.IO/A4DQP
In essence, we estimated the emotional tone of words based on a weighted sample of sentences coded by untrained native speakers on a crowdsourcing platform using a set of carefully translated coding instructions. For each language, a set of 100 sentences was used for testing the crowdcoders’ reliability. To this end, the 100 sentences were coded by two parties: (1) the crowdcoders on the crowdcoding platform Figure Eight and (2) the gold standard team consisting of the first and second author of this article and a bilingual native speaker of the respective language. In the gold standard case, the three persons came up with consensual codings regarding expression of emotion in the 100 sentences. The gold standard and the crowdcoding conform quite well for English, German, and Turkish. For Arabic, the reliability is lower, with raw agreement at .532 (for the data and a methodological reflection, see the online appendix). This low reliability will be discussed later.

After testing the performance of crowdcoders, 3,000 sentences were coded by crowdcoders for each language and randomly separated into a training set (2,400 sentences) and a test set (600 sentences). Using the training set, we calculated the emotional loadings of each unique word in the training set using the method proposed by Haselmayer and Jenny (2017), that is, the number of sentences with that word coded with a specific emotion divided by the total number of sentences containing the word. We determined the optimal size of our dictionary by removing words with emotional loading lower than the cutoff point determined by 10-fold cross-validation. Finally, we determined the criterion validity of our dictionaries using the test set to calculate the area under the receiver operating characteristic curve (AUC). The resulting dictionaries can be found in the online appendix. For example, some sentiment words associated with fear were killer, brutal, violence, and chaos; words associated with pity included condolence, desperate, racism, and tragic.

The customized fear and pity dictionaries have satisfactory AUC, with the exception of the English and Arabic dictionaries for pity, for which the AUC is around 0.48. As studied by Song et al. (2020), suboptimal interrater reliability in human coding can in turn jeopardize criterion validity (e.g., low AUC). We suggest that the low AUC outcomes in the two instances should largely be attributed to poor interrater reliability in the first place. There are several possible explanations for this related to

---

5 Some scholars have voiced legitimate concerns about the ethical issues of crowdcoding (e.g., Williamson, 2016). One suggestion by Williamson is to pay the crowdcoders US$0.5 for a 3-minute task. Few crowdcoding studies have ever reported their pay scheme, but we would like to contribute to illuminating the ethical situation. We paid our crowdcoders at least US$0.03 for coding one sentence; in some languages, this increased to US$0.08. We had ascertained beforehand that an average person can code around five to six sentences per minute. Thus, we paid (0.03*5.5)*3=US$0.495 for a 3-minute task, or US$9.9 per hour. The low end of our pay scheme thus roughly equals the official minimum wage and the remuneration for university-employed student assistants at the time and in the country from which we directed this study (Germany). Depending on the scarcity of their language ability, some crowdcoders clearly earned more than that.

6 There are considerable variations in the size of dictionaries. The number of words in fear dictionaries are 220 (English), 73 (German), 252 (Turkish), and 556 (Arabic). For the pity dictionaries, the numbers are 35 (English), 143 (German), 26 (Turkish), and 650 (Arabic).
crowd-coding: (1) Coding emotions in text essentially amounts to determining a latent variable.\(^7\) The traditional content analysis literature has acknowledged that the coding of latent variables should not be expected to have high intercoder reliability (Neuendorf, 2016). Coder training is essential to ensure interrater reliability, but it is impossible to administer thorough coder training to crowdcoders, and thus there might be interrater variation in what was considered fear or pity.\(^8\) (2) Platform-specific problems such as cheating might have crept into our results, although we have done our best to prevent cheaters (see the online appendix for a detailed discussion).\(^9\) We admit this as a weakness of this and other studies; we suggest that the trade-off between speed and reliability should be critically assessed in each individual study and warrants specific justification (Lind, Gruber, & Boomgaard, 2017). Apart from these reasons, the size of our test set might have artificially reduced the accuracy of our dictionaries because it might not provide enough variation to test the diverse lexicons generated from the training set. Readers should thus be conscious that our dictionaries are preliminary even though they have high face validity.

To answer RQ1 to RQ3, we visualized the variations in emotions in each combination of topics. For each article, the sentiment score of a particular emotion was calculated as the quotient of the frequency of sentiment words of that emotion and the total number of emotion words in all emotions.\(^10\) Then, we aggregated the sentiment scores by calculating the stratified winsorized mean of the scores (i.e., discarding values below the 5th percentile and above the 95th percentile) based on three strata (RQ1: dominant religion; RQ2: ideological position; RQ3: country). A winsorized mean was used instead of the traditional means to control for the influence of extreme values on the mean from the Zipf's distribution of sentiment words in news articles (Piantadosi, 2014). We adjusted the winsorized mean sentiment score from each combination of TMR topics with the counterpart from the controls to

---

\(^7\) Some scholars argue theoretically that discrete categories of latent emotions such as pity and fear do not exist (Barrett, 2009). Empirically, however, there is strong evidence that these discrete categories do exist and can be measured (Cowen & Keltner, 2017). This study follows the previous sentiment dictionary development efforts, such as Moral Foundation Dictionary (MFD) and NRC Word-Emotion Association Lexicon (NRC), by assuming that discrete categories of emotions do indeed exist.

\(^8\) For example, sentences with the Arabic phrase “1948 tragedy” (the Palestine War in 1948) were very likely to be coded as fear or pity by Arabic crowdcoders because they tend to associate these emotions with the event, although the sentences do not explicitly express any emotion.

\(^9\) According to Peer, Brandimarte, Samat, and Acquisti et al. (2017), participants from Figure Eight were found to be more diverse, but more honest, than those from MTurk. The likelihood of validity-threatening cheating should be lower than in other studies using MTurk. However, the diversity in participants could hamper the inter-rater reliability. For the cases of English and Arabic, the diversity of participants is greater than for German and Turkish.

\(^10\) Some English sentences determined to have a high level of fear are: (1) “I still cannot forget the nightmare,” and (2) “Others, like myself, look and see these kidnapped girls, and we think about what would happen to our own children if they were taken from us.” The following are examples of sentences determined to have high levels of pity: (1) “Rawi says: ‘For the first time I felt I couldn’t raise the money I needed in Australia to help the people of Afghanistan’,” and (2) “We mourn those lost to the horrific attacks in Paris.”
calculate the relative risk (RR), that is, the likelihood for an emotion to occur in TMR articles compared with the control articles.

**Results**

In total, 567,925 articles were included in the analysis. The distribution of these articles is presented in Table 3.

<table>
<thead>
<tr>
<th>Country</th>
<th>n</th>
<th>C</th>
<th>R</th>
<th>M</th>
<th>T</th>
<th>MR</th>
<th>TR</th>
<th>TM</th>
<th>TMR</th>
<th>% total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUS</td>
<td>70,866</td>
<td>35,410</td>
<td>3,784</td>
<td>8,907</td>
<td>6,315</td>
<td>1,912</td>
<td>664</td>
<td>8,215</td>
<td>1,659</td>
<td>12.5</td>
</tr>
<tr>
<td>DEU</td>
<td>147,242</td>
<td>79,239</td>
<td>3,7177</td>
<td>9,723</td>
<td>6,035</td>
<td>4,842</td>
<td>2,063</td>
<td>5,716</td>
<td>2,447</td>
<td>25.9</td>
</tr>
<tr>
<td>LBN</td>
<td>68,952</td>
<td>34,270</td>
<td>9,268</td>
<td>10,670</td>
<td>8,406</td>
<td>1,470</td>
<td>1,222</td>
<td>2,872</td>
<td>774</td>
<td>12.1</td>
</tr>
<tr>
<td>CHE</td>
<td>78,081</td>
<td>43,322</td>
<td>1,6448</td>
<td>7,042</td>
<td>3,749</td>
<td>2,009</td>
<td>909</td>
<td>3,498</td>
<td>1,104</td>
<td>13.7</td>
</tr>
<tr>
<td>TUR</td>
<td>66,827</td>
<td>33,407</td>
<td>6,960</td>
<td>7,636</td>
<td>14,834</td>
<td>882</td>
<td>1,272</td>
<td>1,442</td>
<td>394</td>
<td>11.8</td>
</tr>
<tr>
<td>USA</td>
<td>135,957</td>
<td>67,984</td>
<td>7,735</td>
<td>2,1403</td>
<td>14,338</td>
<td>3,188</td>
<td>1,479</td>
<td>15,551</td>
<td>4,279</td>
<td>23.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% total</th>
<th>51.7</th>
<th>15.0</th>
<th>11.5</th>
<th>9.45</th>
<th>2.52</th>
<th>1.34</th>
<th>6.57</th>
<th>1.88</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>% total</td>
<td>51.7</td>
<td>15.0</td>
<td>11.5</td>
<td>9.45</td>
<td>2.52</td>
<td>1.34</td>
<td>6.57</td>
<td>1.88</td>
<td>100</td>
</tr>
<tr>
<td>% total</td>
<td>31.1</td>
<td>23.8</td>
<td>19.6</td>
<td>5.2</td>
<td>2.8</td>
<td>13.6</td>
<td>3.9</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>% total</td>
<td>31.1</td>
<td>23.8</td>
<td>19.6</td>
<td>5.2</td>
<td>2.8</td>
<td>13.6</td>
<td>3.9</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>% total</td>
<td>20.5</td>
<td>10.9</td>
<td>53.4</td>
<td>15.3</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. C = controls; AUS = Australia; DEU = Germany; LBN = Lebanon; CHE = Switzerland; TUR= Turkey.

Germany and the U.S. contributed almost half of the articles, and the most common topic was refugees alone (15%). Articles with combined topics contributed around a quarter of all articles mentioning T, M, or R. Among these articles with combined topics, half combined terrorism and Muslims. About 12% of the articles combined refugees with other topics.

Figure 2 shows the RRs for all topical categories. It is striking that the distribution of fear across topical categories (left side of Figure 2) is an almost exact mirror image of the distribution for pity (right side). When fear is stronger in a topical category, pity is weaker in that same category, and vice versa. This “butterfly” pattern shows that fear and pity generally work antagonistically. In addition, refugee coverage exhibits the lowest level of fear and the highest incidence of pity, whereas the opposite is true for terror-related topics, which feature much fear and relatively little pity. Across all countries and outlets, therefore, we confirm the patterns reported in earlier studies. But what happens when we break our data down to answer our research questions?
Figure 2. Variations in fear and pity by topic. Note. RR = relative risk; a higher value indicates higher likelihood to occur than controls, and vice versa.

**RQ1: Comparison of Fear and Pity by Dominant Religion**

Based on the majority religion (Muslim majority in Turkey and Lebanon, Christian majority in all other countries), the RR values were calculated for each combined topic (Figure 3).

First, we observe the same butterfly pattern in both country groups, although the pattern in Christian-majority countries resembles the pattern in Figure 2 more closely. Second, when we look at the two country groups separately, we observe that in Christian-majority countries, fear increases and pity decreases when the topic of Muslims is combined with terrorism (M versus TM). Interestingly, the addition of the refugee topic into the mix decreases the expression of fear and increases pity in Christian-majority countries (M vs. MR, T vs. TR, TM vs. TMR). In Muslim-majority countries, combining the topic of Muslims with other topics does not boost fear or reduce pity in most of the cases, with one exception: a slight increase in fear when comparing M and MR. This analysis answers RQ1 and supports the influence of the majority religion in a country on the emotional tone of coverage about Muslims and combined topics involving Muslims.
**Figure 3.** Religion-specific variations in fear and pity by topic. Note. RR = relative risk; a higher value indicates higher likelihood to occur than controls, and vice versa.

**RQ2: Comparison of Fear and Pity by Political Alignment of Outlets**

When stratified by political alignment of the media outlets, we again find the butterfly pattern between the distributions of fear and pity across topical categories (Figure 4). In addition, both emotions show relatively little variation across outlets with different political leanings.
Figure 4. Ideology-specific variations in fear and pity by topic. Note. RR = relative risk; a higher value indicates higher likelihood to occur than controls, and vice versa.

The patterns of variation in fear when two or three topics are combined are similar across the three political alignments. The only remarkable contrast is in the magnitude of such differences—for example, the combination of T and M generates higher increases of fear (over M and T alone) among right-wing media than among center and left-wing media. Relatively speaking, political alignments do not exert a prominent influence on the changes in pity and fear when topics are combined.
Figure 5 displays the prevalence of fear and pity in each of the six countries separately.

Figure 5. Country-specific variations in fear (top) and pity (bottom) by topic. Note. RR = relative risk; a higher value indicates higher likelihood to occur than controls, and vice versa.
One important feature of the figure is that the English pair (USA/Australia) and the German pair (German/Switzerland), respectively, have similar ranges in the prevalence of emotions. This can be an indication of residual influence of language differences on the results. Even with these residual influences, there are still considerable variations in how some topics are reported between countries of the same language group. For example, R-only and MR topics are reported with more fear and less pity in the U.S. than in Australia. A similar pattern is also observed when comparing Germany with Switzerland. This hints to the fact that peculiarities of language use alone cannot explain country differences in emotional tone (even though it does play a role), but that idiosyncratic national circumstances intervene.

Such idiosyncrasies can be best demonstrated with the pronounced difference between Turkey and the other countries. Turkish online news media exhibit more fear in terrorism-only and TR articles than in control articles, but less fear than in controls for all other topics. For pity, the pattern is even more pronounced: T-only, TM, and TR articles feature much more pity than articles on other topics. A likely cause is that in Turkish government communication, anything related to the Kurds is labeled terrorist or support for terrorism. The higher fear and pity values for T and TR are thus not necessarily always related to actual terrorist attacks. Although there were some high-profile Islamist attacks in Turkey (e.g., Ankara bombing) during the study period, most of the Turkish coverage mentioning terrorism was about the Kurdish issue. Therefore, we observed a unique emotional profile that sets Turkey apart from the other countries studied.

In addition, the Lebanese online news sources surprisingly exhibit more fear and less pity in refugee-related articles than news sources in the Western countries; in Turkey, this holds for pity, but not for fear, which is very low in Turkish refugee articles. This shows that cultural proximity to the refugee cohorts alone does not explain the emotional tone in refugee-related coverage—an indication that social and political circumstances and/or journalistic conventions prevalent in individual countries do account for at least part of the variation.

To further explore country-specific circumstances, we conducted a supplementary analysis of pity and fear of R and MR articles for all German-language outlets before and after New Year’s Eve 2015/2016 (Figure 6). It is generally believed that the Kölner Silvesternacht incident changed the discourse about refugees in Europe, especially in Germany, but possibly also neighboring countries such as Switzerland. Comparing pity and fear before and after the incident, almost all outlets displayed a reduction in pity and an increase in fear after the event in their refugee-related coverage. This case study confirms again that fear and pity are largely antagonistic to each other in the news language not only across topics, but also across time. At the same time, we observed different changes before and after the incident: Swiss media covered R and MR issues with slightly lower fear and greater pity before the incident than German media. Moreover, Swiss media also demonstrated slightly larger jumps after the incident.

---

11 Exceptions can be found in German public broadcaster ARD’s news website Tagesschau, the liberal weekly Die Zeit, and the Swiss daily Tagesanzeiger, which warrants further investigation.
Figure 6. Variations in pity and fear of R and MR articles before and after Kölner Silvesternacht among German-language outlets. Note. RR = relative risk; a higher value indicates higher likelihood to occur than controls, and vice versa.

To sum up, we have observed considerable variations in pity and fear across the six countries. These variations can be explained by a combination of residual language differences, idiosyncratic national circumstances, and differential responses to events that are significant to a particular national or language community.

Discussion

In this article, we used a cross-cultural comparative design to study fear and pity in more than 560,000 articles from 37 online news outlets in six countries regarding the topics of terrorism, Muslims, and refugees (TMR). The design includes both Christian-majority (U.S., Australia, Germany, and Switzerland) and Muslim-majority (Turkey and Lebanon) countries, and therefore, our study fills a research gap in the literature. Using a multilingual set of human-validated sentiment dictionaries, the current study compared the emotions in TMR coverage and studied how emotions vary when the topics
are combined. When we studied multiple media outlets together and adjusted them using a group-specific baseline, we identified both commonalities and discrepancies in the emotional profile of TMR coverage.

On the commonalities, we identified the butterfly pattern of fear and pity (i.e., when one emotion goes down in a topical category, the other goes up) across all countries and outlets (Figure 2), but also when we broke the data up by majority religion (Figure 3) or by the outlet’s political leaning (Figure 4). The butterfly pattern shows that fear and pity do work antagonistically in the news language quite universally. When looking at relative differences among topics, journalists seem to be reporting TMR topics using a set of transcultural rules: lowest fear and highest pity for refugee-only articles, and reduced pity and increased fear when comparing M and MR. However, mixing R into terrorism-related topics can usually increase pity and reduce fear. These universal patterns have not been demonstrated empirically before. Our findings fit into Chouliaraki’s (2006) model of media discourse on human suffering, which holds that refugee coverage is more likely to be of “ecstatic” nature, and journalists can have more room to express pity.

Although we observed the transcultural butterfly pattern, we also found discrepancies across individual countries (Figure 5). Three factors can explain these discrepancies: residual language differences, idiosyncratic national circumstances (e.g., Turkish use of “terrorism”), and differential responses to TMR-related events. Although the measurement-related problem of residual language differences might explain some of the variation in Figure 5, the latter two country-level factors also accounted for a considerable amount of variation. Using our butterfly analogy, we can observe different species of butterfly in various countries, probably because these species (emotional profile of TMR coverage) survive better in their respective “natural habitats” (i.e., the national circumstances). The additional analysis of the Kölner Silvesternacht (Figure 6), which has naturally adjusted for the measurement problem of language differences, demonstrated how we should reconcile the similarities in pity and fear across topics (the general trend of increasing fear and decreasing pity) and the discrepancies across nations (Switzerland displayed a different baseline level and a different slope of change than Germany). Our findings speak against the notion of disappearing national public spheres and a general global homogenization of news content (Wessler, Peters, Brüggemann, Kleinen-von Königslöw, & Sifft, 2008). Instead, we reconfirmed that the country in which a media outlet operates still plays an important role in shaping TMR coverage (e.g., Gerhards & Schäfer, 2013; Ruigrok & van Atteveldt, 2007).

We contend that previous findings based on only one culture should not be considered globally applicable. For example, although previous studies found that certain right-leaning media report TMR issues in a more hostile and exclusionary way than other media (Aday, 2010; Aday et al., 2005; Berry et al., 2016), our cross-cultural data suggest that political alignment of media outlets plays a relatively small role in explaining variation in emotions (our RQ2). Thus, the role of political alignment identified by previous studies might only be a local phenomenon (presumably in Western countries) rather than a global pattern. Having said that, we acknowledge that we cannot assess the measurement invariance of our political alignment coding (“Is the meaning of center-left the same in the U.S. and Turkey?”). However, we believe that the apparently large difference in explanatory power between national-level variables and outlet-level political alignment cannot be explained by the measurement invariance of our political alignment variable alone. A slight improvement in the political alignment classification should thus
not generate a large difference. Nonetheless, we advocate for the development of an internationally valid methodology to evaluate the political alignment of media outlets. Other country-level independent variables, such as the ideological leaning of a country’s government and the presence and relevance of populist parties, might be useful to include in this analysis too.

When we turn to practical lessons that can be drawn from our results, it is important to note that the reporting on refugees alone was usually not negative (lowest in fear, highest in pity). However, when the discourse of refugees was mixed with one or both of the other two topics, the fear tone was more prominent, and pity was reduced in coverage of media outlets in Western countries. This high-fear/low-pity tone might cast negative out-group perceptions of refugees in general and hinder their social integration. However, journalists also need to inform the public about the reality of a crisis. Journalists might face a dilemma when reporting the Muslim background of refugees or about incidents in which the perpetrators of radical Islamist terrorist attacks had managed to acquire refugee status in their target country beforehand. Similar to the earlier recommendation on news differentiation between Muslims and Islamist terrorists (von Sikorski et al., 2017) and the call for responsible reporting on Muslims (K. Neumann et al., 2018), we call for responsible reporting on refugees as a precautionary measure. Avoiding unnecessary generalizations, newsmakers should clearly separate the majority of refugees from deviant individuals who abuse their refugee status and support radical Islamist terrorism.

Although our study design fills the cross-cultural research gaps and enhances the complexity of our knowledge regarding the issues at hand, it also has its share of limitations. First, our definitions of T, M, R topics are based on keywords and thus cannot determine the targets of the sentiment expressed in the article. For example, the article-level sentiment analysis cannot distinguish whether the pity expressed in a TMR article is targeting the victims of a terrorist attack or refugees. To rectify this limitation, we plan to extend the current study using a semantic network approach to study how emotions are targeted toward different actors (van Atteveldt, 2008).

Second, the country differences we observe (Figure 5) might partly be due to the residual language differences in the performance of our sentiment dictionaries. However, it is difficult to tease out the country differences from language differences because they are correlated. One option is to further improve the sentiment analysis so that sentiment detection is equally good across all languages. To this end, machine learning methods have been suggested previously (González-Bailón & Paltoglou, 2015), while traditional human coding might even be a better choice (e.g., in Gerhards & Schäfer, 2013), but prohibitively expensive in big data contexts. In any case, improvements in multilingual, cross-cultural sentiment analysis are welcome and needed.

To be sure, we advocate for the replication of our findings because more reliable approaches to coding emotions in text might become available in the future. More generally, too, we call for systematic tests of the mechanisms identified in our research questions on specific emotions to arrive at more targeted, situated explanations of emotional tone in the news. We also hope that our study encourages more cross-cultural research into the emotional implications of issue combinations, particularly in cases such as ours, in which emotions expressed in the media are directly related to intergroup conflict and peace.
References


