



**UvA-DARE (Digital Academic Repository)**

**Multimodal news framing effects**

Powell, T.E.

[Link to publication](#)

*Creative Commons License (see <https://creativecommons.org/use-remix/cc-licenses/>):*  
**Other**

*Citation for published version (APA):*  
Powell, T. E. (2017). *Multimodal news framing effects*.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## Chapter 2

# A clearer picture: The contribution of visuals and text to framing effects

---

This study is published as: Powell, T. E., Boomgaarden, H. G., De Swert, K., & de Vreese, C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. *Journal of Communication*, 65(6), 997-1017

---

### **Abstract**

Visuals in news media help define, or frame issues, but less is known about how they influence opinions and behaviour. I use an experiment to present visual and textual exemplars of frames from war and conflict news in isolation and in image-text congruent and incongruent pairs. Results show that, when presented alone, images generate stronger framing effects on opinions and behavioural intentions than text. When images and text are presented together, as in a typical news report, the frame carried by the text influences opinions regardless of the accompanying image, whereas the frame carried by the image drives behavioural intentions irrespective of the linked text. These effects are explained by the salience enhancing and emotional consequences of images

---

Graphic visuals are a prominent part of today's news coverage but how they affect public opinion and behaviour remains unclear. This is important since emotionally charged images are a powerful vehicle for the framing of political messages – conflicts in Libya, Palestine, Syria and Iraq provide recent examples – and emotional responses influence the decisions of public and politician alike (Damasio, 1996). In this study I presented visual and textual exemplars of frames from war and conflict news to shed new light on how visuals contribute to framing effects. In doing so I position visuals within a nascent *multimodal* verbal and visual framing theory.

News frames are “interpretive packages”, consisting of a “central organising idea”, that are used by journalists to communicate the most salient aspects of an otherwise complex story to a reader (Entman, 1993; Gamson & Modigliani, 1989, p.3). By making these aspects more applicable in the mind of the reader, frames affect *how* people think about issues (Scheufele & Tewksbury, 2007). This has been shown to influence readers' information processing and recall (e.g., Newhagen & Reeves, 1992), attitude formation (e.g., Nelson et al., 1997), and opinions and behavioural intentions (e.g., Schuck & de Vreese, 2006). However, with a few exceptions (e.g., Arpan et al., 2006; Gibson & Zillmann, 2000; Iyer, Webster, Hornsey, & Vanman, 2014; Pfau et al., 2006), this research has focused on framing effects produced by an news story's text, while the effect that *images* have on the viewer remains relatively under-researched (e.g., Coleman, 2010).

News visuals fulfil a unique role in international affairs by shedding light on far flung conflicts and bringing unfamiliar and complex stories to life. Indeed, graphic images of militiamen and their slain victims are beamed across the globe to meet the aesthetic demands of 21<sup>st</sup> century news production standards. Furthermore, a shift towards contemporary formats in which eye-catching images are presented with accompanying text kept to a minimum means that an image's impact when viewed alone is no longer a trivial matter. Of course, news images still typically appear in the context of a report and therefore should also be considered alongside an accompanying text. Assimilation of these elements into a story, however, is not always done by the journalist themselves, or image selection can be from a limited set provided by wire services and based on a superficial notion of what might have an ‘impact’ on the reader (Huang & Fahmy, 2013). Whatever the cause, haphazardly paired images and text have been shown to open the door to unexpected and as yet unquantified effects (Fahmy et al., 2014).

Framing effects of images and texts depend on their unique characteristics (Geise & Baden, 2014). Images serve to index and reproduce reality (Messaris & Abrahams, 2001) and are attention-grabbing (Garcia & Stark, 1991). By evoking a heightened emotional experience compared to text (Iyer & Oldmeadow, 2006), visuals connect with a reader and in turn might also be more persuasive. In contrast, text is less salient, but possesses a clear structure for inferring who did what to whom and why (Entman,

1993). Clarification of the individual and combined effects of images and text would be an important insight for those in the newsroom, and a first step towards integrating visuals into a multimodal framing theory.

This study used an experimental manipulation of an online news report to pursue this goal. I presented images of victims and militants from a little known conflict and corresponding text frames in isolation and in congruent and incongruent image-text pairs. I measured how feelings of sympathy, fear and anger influenced participants' support for military intervention and behavioural intentions towards the conflict. In doing so, I aimed to address three questions: Firstly, do images generate stronger framing effects than text when viewed in isolation? Secondly, what are the combined framing effects when image and text are presented together in different configurations? Finally, what role do emotionally charged visuals play in producing these effects?

## The power of visuals

Scholars have long acknowledged visuals as a vehicle for news frames by visualizing and emphasizing a particular aspect of an issue (e.g., Grabe & Bucy, 2009; Graber, 1990). Plentiful examples come from war and conflict news (e.g., Fahmy & Neumann, 2011; Greenwood & Jenkins, 2013; Schwalbe, 2006), with one instance being coverage of the fall of Srebrenica in 1995. In this case, graphic images of the massacre of thousands of Bosniaks contributed to what Robinson (2002, p.78) called "empathy framing". These images highlighted the *obligation* of the international community to intervene to prevent human rights violations, a topic of continued relevance since the end of the Cold War (Evans, 2004). In contrast, the *risks* of intervening – emphasised by images of belligerent militants – had, until the massacre, provided a strong counter-frame (Entman, 2003), and one that has also been prominent in coverage of recent conflicts in Libya, Syria and Iraq. Based in part on the lessons from Bosnia, Robinson (2002) proposed a model of media effects, stating that empathy-driven frames that are also critical of an indecisive government can influence policymaking.

Despite the intuitive link between powerful images and policy support (Domke et al., 2002), empirical knowledge of visual framing effects lags behind text in an increasingly multi-modal media reality (Coleman, 2010). Such insights are complicated by the multi-level "winnowing process" that determines how visuals contribute to frames (Schwalbe, 2006, p. 269). First, those events deemed newsworthy are 'framed' by the photojournalist's selection of what to capture and how (e.g., angle, perspective, cropping). In the newsroom, editors choose which images to publish beside an article's text, and in what size and position. These stylistic-semiotic decisions have a connotative influence over how visuals are decoded by the reader (Kress & van Leeuwen, 1996) and how they resonate with the viewer's internal frames of reference. In this study I employ obligation and risk frames in online news content<sup>1</sup> to explore the final level of this

process – how images interact with text to influence political opinions and behaviours.

In contrast to their political effects, the impact of visuals on basic psychological processes is well established. For instance, the meaning of images is accessed faster than that of text (Barry, 1997). In the realm of memory and learning, the picture superiority effect describes how named images are better recalled than named words due to the concrete imagery they generate (Paivio, 1991; Nelson et al., 1976). These findings have been extended in the context of media and communication, where news visuals have been shown to be more memorable, leading to better recall of images over text (e.g., Newhagen & Reeves, 1992).

Beyond memory, images are also more attention-grabbing than text. Eye tracking studies have shown that images are the most common entry point into newspaper pages (Garcia & Stark, 1991), which points to their amplifying effect on psychological processes. In this study I presented images-alone, text-alone, and text accompanied by an image, and confirmed that stimuli with an image were rated as more attention-grabbing than stimuli without.<sup>2</sup> Inclusion of an image, then, can increase the salience of one issue compared to the another, which in turn can raise it up the public agenda (Wanta, 1988) and prime these topics in the minds of readers (Scheufele & Tewksbury, 2007). What remains unclear, however, is the link between salience and framing effects: Are visuals just eye-catching, or do they play a more meaningful role in highlighting particular parts of a message and influencing the reader? This is the principal question of visual framing effects.

## Visual framing effects

Social psychologists studying attitude change made early inroads to clarify the persuasive effects of visual messages. A classic study by Chaiken and Eagly (1976) showed that (audio-)visuals were most effective at changing opinions when a message was easy to understand, whereas written text was more effective when a message was complex. Chaiken (1980) used this and related insights to develop a dual-route theory of persuasion - the *heuristic-systematic model*. She distinguished between systematic processing, an analytic orientation in which the recipient scrutinizes an argument, and heuristic processing, in which inferential rules based on a subset of information are used to form opinions. Furthermore, Sparks et al. (1998) argued that visuals should foster relatively more heuristic versus systematic processing, and vice-versa for text. These predictions have garnered some support in marketing psychology literature (e.g., Sojka & Giese, 2006) but have received little scrutiny in media effects research despite the continued relevance of dual-route models of persuasion (e.g., Petty & Cacioppo 1986).

A small but rapidly expanding body of experimental research has since reinvigorated interest in visual framing effects. Two experiments by Gibson, Zillmann and colleagues

showed that manipulation of an image whilst keeping the accompanying text constant can influence perceptions of an issue and later recall of the image's content (Gibson & Zillmann, 2000; Gibson, Zillmann & Sargent, 1999). Arpan et al. (2006) used images of social protests to show that images depicting higher levels of conflict led to more negative evaluations of the protest and protestors, but only when the issue was of interest to participants. More recent studies focused on international affairs to show that military images, particularly graphic photographs (Scharrer & Blackburn, In-Press) and those showing conventional images of loss (Gartner, 2011) can decrease support for war. Taken together, these studies point towards picture superiority effects. In contrast, Pfau et al. (2006) showed that the text accompanying an image can also play a role. They showed that images of the Iraq war with a caption, compared to images with full-text and text presented alone, led to a small decrease in support for U.S. presence in Iraq and an increase in negative emotional responses. Domke et al. (2002) criticised the widely held notion that vivid images drive public opinion. They argued that the inclusion of an image influences opinions, but the specific content of an image interacts with participants' existing knowledge to shape information processing and judgements.

Geise and Baden (2014) assimilated these insights and contrasted the characteristics of images and text to formulate a theoretical framework for *multimodal* visual and verbal (textual) framing effects. At the heart of their argument is that visuals have an "amplifying effect", are perceived quickly and are highly salient, and their analogical quality can make them easy to understand. However, depending on an image's compositional components and the prior knowledge of the recipient, various interpretations and effects can result (especially for highly ambiguous images). In contrast, text is typically less salient, with framing primarily achieved through the inclusion or omission of selected information which is fine-tuned thereafter (Geise and Baden, 2014). Text also has a conventional and highly learned syntax, which allows a more prescribed construction of meaning. However, if the recipient lacks required knowledge, of both of the language and the context, the text will not be decoded and framing will fail (Scheufele & Tewksbury, 2007).

To our knowledge, no studies have explicitly tested these predictions about the ability of images alone to drive news framing effects, nor of the relative power of images alone compared to text. Based on the unique aspects of text and images, and particularly the power of images, I hypothesised that:

*H1: When viewed in isolation, images and text both generate framing effects, with images being a stronger influence than text.*

## Multimodal framing effects

While the individual contribution of images and text to framing effects is interesting, their interactive effects is an underexplored and less predictable phenomenon relevant to modern media practice (e.g., Fahmy, Bock, & Wanta, 2014).

When media messages are presented in both visual and verbal modalities, learning and memory is improved (Paivio, 1991). These effects might be further enhanced when these messages are highly *congruent*. By congruence, I refer to whether the text and image (or, in the case of television news, the audio and visual) are matching or not. A handful of empirical studies have used TV news to show that the addition of a congruent visual input to an audio input improved memory and issue understanding (Graber, 1990; Reese, 1984), with messages in separate modalities fusing over time (e.g., Drew and Grimes, 1987). However, when effects of congruent stimuli have been compared with incongruent results have been mixed, with a tendency for improved memory for the visual stream (e.g. Lang, 1995). At best, these findings suggest that when compared to incongruent stimuli, congruence increases media learning and salience. However, the impact of different combinations of multimodal messages on opinions and behavioural intentions is not known.

Congruent multimodal messages should be endowed with the salience, vividness and memorability of visuals, plus the guided structuring of text (Geise & Baden, 2014). When incongruent, the body of research highlighting the power of visuals suggests that attention-grabbing images may dominate and override divergent textual information (Gibson & Zillmann, 2000). On the other hand, it is plausible that the explicit structure of text could drive framing effects by disambiguating an image's content and guiding interpretation (Huang & Fahmy, 2013). Indeed, instead of exerting a persuasive influence, the vivid qualities of images have been shown to merely attract attention and enhance the salience and recall of a linked text (Zillmann, Knobloch & Yu, 2001).

This study used two opposing media frames to shed new light on the relative impact of images and text when combined in congruent and incongruent pairs. Given the reported power of images and positive effects of congruence, but considering the uncertainty over whether image or text exert a defining influence, I tentatively proposed the following hypotheses:

*H2a: Framing effects are strongest when image and text are congruent, compared to incongruent.*

*H2b: When image and text are incongruent, images are more important in determining framing effects.*

## Emotions as a mechanism for visual framing effects

Time magazine's Lance Morrow (1993) argued that powerful news images from foreign conflicts "are mainlined directly into the democracy's emotional bloodstream without the mediation of conscious thought" (p. 36). Quantification of this claim comes from psychology and neuroscience research which shows that, in concert with more 'rational' processes, emotions play an important role in heuristics (mental rules of thumb) that are indispensable for human decision making (e.g., Damasio, 1996). This is particularly so when available information is incomplete or complex (Slovic et al., 2007). Since the very function of news frames is to reduce complex issues into interpretative packages, it follows that emotions elicited by powerful visuals could play a central role in framing effects.

Although a large body of non-visual research has shown that emotions can shape individuals' political behaviour (e.g., Neuman, Marcus, Crigler & Mackuen, 2007), few studies have assessed the contribution of affective responses to visuals in the framing process (Brantner, Lobinger, & Wetzstein, 2011; Iyer et al., 2014; Pfau et al., 2006). This is particularly surprising given that visuals play a leading role compared to text in triggering an emotional response (Iyer & Oldmeadow, 2006).

One area in which the emotional consequences of visuals have been credited is agenda setting research (Coleman, 2010; Wanta, 1988). For instance, Fahmy et al. (2006) used the context of 9/11 to show that memories of visuals from the attacks, and participants' emotions, particularly sorrow and shock, predicted their sample's concern with terrorism. As such, emotionally charged visuals can increase an issue's salience, but less is known about how visuals evoke emotions that, in turn, influence opinions and behaviours towards issues.

The mediating role of emotions in visual framing effects has been examined by Iyer et al. (2014). They showed participants a neutral text describing the 2005 London bombings followed by photographs, either focusing on victims or terrorists. Their results conveyed that the content of images elicited specific emotions of sympathy, fear or anger, which in turn predicted support for specific policies. In this way, specific "emotivational" goals expressed via an individual's emotional response to an image can give rise to distinct attitudes and behavioural intentions (Frijda, 1988).

In this study, I measured how participants' sympathy, fear and anger in response to image and text frames emphasising the obligation to protect suffering victims, versus the risks posed by belligerents, mediate support for military intervention and behavioural intentions towards a foreign conflict. For example, a frame conveying suffering victims should elicit sympathy, and sympathy should encourage helping behaviour (Iyer et al., 2014). A frame emphasising dangerous militants should evoke fear, and provoke an avoidance response (Lazarus, 1991). Anger, triggered by a negative situation where responsibility can be attributed to an actor (in this case the belligerents in the conflict) should prompt an individual to act to confront the problem (Frijda, 1988). Crucially,



our experimental manipulations should clarify the specific role of images and text in delivering these effects.

Given the heightened potential of visuals to evoke an emotional response and the influence emotions can have on our decision making, I proposed the following hypothesis:

*H3: Indirect effects of emotions on framing effects are stronger for images than text for both image/text-alone conditions and image-text combinations.*

## Method

### Design

To test these hypotheses I conducted an online survey-embedded experiment. Participants were randomly assigned<sup>3</sup> to the following conditions: For the image-text combinations, six conditions were arranged in a two (image frame: obligation, risk) by three (text frame: obligation, risk, control) factorial design.<sup>4</sup> Two image-only conditions (obligation, risk) were used to isolate framing effects of images alone. Three text-only conditions (obligation, risk, control) were used to measure framing effects of text alone.

### Participants

A total of 1233 American adults aged 18-65 were recruited via Research Now, a data collection company. 128 of these were excluded because they dropped out before completing two-thirds of the survey. 12 were excluded for straight-lining behaviour (answering all questions with the same response). 11 who reported that they did not see the stimulus properly were also removed. The final sample included 1082 participants. The median time taken to complete the survey was 10.5 minutes, and additional checks suggested that participants were attentive during the experiment.

The sample was representative of the U.S. population (United States Census 2010) for age ( $M = 40.6$ ,  $SD = 13.24$ ) and sex (577 females). Participants came from a variety of racial backgrounds, although most were white and US-born. Participants came from a range of educational backgrounds, and on a continuous measure of political orientation were almost perfectly normally distributed around 'moderate' (1 = *liberal*, 7 = *conservative*;  $M = 4.0$ ,  $SD = 1.55$ ). More details about the attentiveness checks and participant demographics can be found in *Appendix A1*.

### Stimuli

Stimuli were selected from media coverage of the ongoing conflict in the Central African Republic (CAR). The conflict in CAR was chosen since it has received little

media attention, therefore reducing the likelihood that prior exposure would influence participants' responses.

Two pilot experiments were conducted to choose stimulus images and articles that conveyed the respective frames and, importantly, to control for several factors that have previously been shown to influence framing effects (e.g., Schuck & de Vreese, 2006). The chosen images and articles were matched for perceived arousal, valence, salience, ambiguity, complexity, and credibility. The images (obligation image from *The Guardian, UK*; risk image from *L'Express, France*) clearly depicted victims of the conflict in the obligation condition and belligerent militants in the risk condition and were matched for size (130 cm<sup>2</sup>). The selected articles were originally downloaded from the BBC website and were shortened and modified to achieve the framing conditions. The obligation and risk versions were matched for length (277 and 278 words, respectively), whilst the control version was necessarily shorter (209 words) due to the removal of words and phrases used to achieve the frame manipulations.

Before combining the images and text in congruent and incongruent pairs, an additional pilot experiment was conducted to maximise congruence between image and text. An example incongruent stimulus combining a risk article with an obligation image, and the rationale for stimulus selection with details of the three pilot experiments can be found in *Appendix A2* and *Appendix A3*.

## Procedure

Upon entering the survey, participants were asked to respond to questions that measured their knowledge of the conflict in CAR and their political orientation.

Participants were then randomised to one of the stimulus conditions, which were presented on a blank screen. In the image-text combination and text-only conditions, participants were forced to spend at least 30 seconds viewing the stimulus. The pilot experiments showed this was long enough to ensure participants had enough time to process the stimuli and short enough that fast readers were not frustrated. In the image-only conditions, images appeared for 10 seconds after which the survey automatically progressed. Immediately before the stimulus, participants were informed that they would be viewing a news image/article about the conflict in CAR and clearly told about the time they were given to do so.

Next the dependent measures were displayed. Measures of policy support, behavioural intentions, salience, emotions and manipulation checks were shown on successive pages. At the end of the survey participants were asked to provide basic personal information, including age, sex, education level, race and birth country.

## Measures

Participant's knowledge of the conflict in CAR was measured prior to stimulus exposure (1 = *Not at all knowledgeable*, 7 = *Extremely knowledgeable*). Geise and Baden (2014) highlighted the potential for knowledge to influence visual and textual framing effects, a finding that is well established in the political communication literature (e.g., Nelson et al., 1997).<sup>5</sup> As expected, the present sample was relatively unaware of the conflict ( $M = 2.76$ ,  $SD = 1.55$ ), but there was high variability ranging from not at all to extremely knowledgeable.

After viewing the stimulus, participants answered questions that measured the key dependent variables. Support for the policy of military intervention in CAR was measured with three items (e.g., "*Sending an international peacekeeping force*"; 1 = *strongly oppose*, 7 = *strongly support*;  $\alpha = .84$ ). Behavioural intentions were measured using four items (*intention to discuss, donate, sign a petition, protest*; 1 = *very unlikely*, 7 = *very likely*;  $\alpha = .87$ ). Appraised salience of the stimulus was measured using one measure ("*The image/article was attention-grabbing*"; 1 = *not at all*, 7 = *strongly*). Finally, participants indicated the extent to which they felt various emotions while viewing the stimulus (1 = *not at all*, 7 = *extremely*), using the same measures as Iyer et al. (2014). Three items assessed fear (*afraid, anxious, frightened*;  $\alpha = .88$ ), three items measured sympathy (*sympathetic, compassionate, empathetic*;  $\alpha = .90$ ), and three items assessed anger (*angry, outraged, furious*;  $\alpha = .91$ ). A measure of distress was also included, in order to rule out the possibility that the stimuli in one condition were more disturbing or more generally emotionally moving than the other. Three items assessed distress (*worried, upset, anguished*;  $\alpha = .87$ ).

## Manipulation Checks

To check that the framing manipulations worked, two questions assessed the extent to which the stimuli conveyed the obligation to intervene in the conflict in CAR (1 = *strongly disagree*, 7 = *strongly agree*;  $r = .62$ ,  $p < .001$ ), and two measures assessed the extent to which the stimuli conveyed the risks of intervening in the conflict ( $r = .78$ ,  $p < .001$ ). A "frame index" was then calculated by subtracting each participant's score for the risk scale from their score on the obligation scale, to indicate the extent to which the stimuli conveyed obligation versus risk. Importantly, there was a significant difference in frame index score between the obligation and risk stimuli for both images and text (both  $p < .001$ ), and the control text achieved a good balance between the two. These results confirm findings from the pre-tests (see *Appendix A3*) that the frame manipulations were effective.

## Analysis

A series of ANOVAs were conducted to examine mean differences between the frame conditions for the dependent variables of support for intervention and behavioural intentions. The indirect effect of emotions in generating framing effects was assessed via ordinary least squares path analysis using Hayes PROCESS-macro in SPSS (Hayes, 2013). For all mediation models the obligation and risk frame conditions were included as the two levels of the independent variable. Participants' reported sympathy, fear and anger whilst viewing the stimulus were included in parallel as mediators. The indirect effect of each emotion on support for intervention and behavioural intentions was examined while controlling for the other emotions. 95% bias-corrected bootstrap confidence intervals based on 10,000 bootstrap samples were used for statistical inference of indirect effects.

## Results

### Images alone vs. text alone

Stronger framing effects were observed in the image-alone conditions compared to the text-alone conditions (see *Table 1*).

For the image-alone conditions, participants reported higher support for intervention in the CAR after they viewed an obligation image compared to a risk image,  $F(1, 228) = 3.76, p = .05, \eta_p^2 = .02$ . The same result was observed for participants' reported behavioural intentions,  $F(1, 228) = 9.74, p = .03, \eta_p^2 = .02$ . For text-alone, the differences between the frame conditions for support for intervention,  $F(2, 167) = 0.24, p = .79$ , and for behavioural intentions,  $F(2, 167) = 1.31, p = .27$ , were non-significant.

Small and non-significant framing effects in the text-alone conditions were counter to our expectations and were explored further via a moderation analysis (Hayes, 2013). The effect of the text frames on support for intervention was moderated by participants' issue-specific knowledge,  $R^2 \text{ change} = .05, F(1, 109) = 5.87, p = .02$ . Probing the interaction (see *Figure 2.1*) showed that only participants with high knowledge of the conflict in CAR, and not those with low knowledge, were significantly more supportive after reading an obligation text compared to a risk text,  $MDiff = -.60, t = -2.04, p < .05$ . This shows that, when presented alone, the text frames did in fact affect participants' support for intervention, but only for those with high issue-specific knowledge. This is examined further in the Discussion.

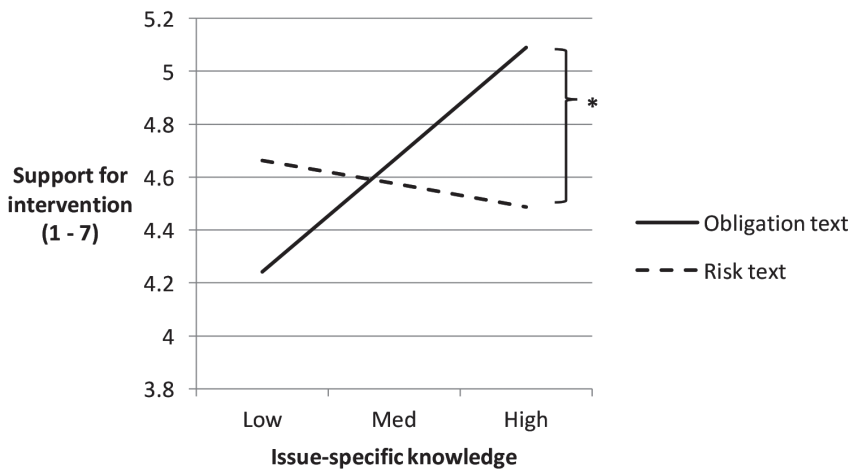
These results support hypothesis 1: *framing effects produced by images alone were stronger than by text alone*. This was further supported by additional analyses that showed framing effects were statistically significantly stronger for images than text for the behavioural intentions variable.<sup>6</sup>

Table 1: Mean support for intervention and behavioural intention ratings per frame condition.

	Image alone		Text alone		
	Obligation	Risk	Obligation	Control	Risk
Support for intervention	4.43* (1.17)	4.09* (1.42)	4.66 (1.31)	4.51 (1.19)	4.57 (0.91)
Behavioural intentions	3.16* (1.43)	2.75* (1.38)	2.92 (1.51)	2.75 (1.30)	3.17 (1.25)
N	113	117	59	57	54

Note. Standard deviations in parenthesis. Means with matching superscripts are statistically different,  $p \leq .05$ . Support for intervention scale: 1 = *strongly oppose* to 7 = *strongly support*. Behavioural intentions scale: 1 = *very unlikely* to 7 = *very likely*.

Figure 2.1. Moderation of text framing effects on support for intervention by issue-specific knowledge.

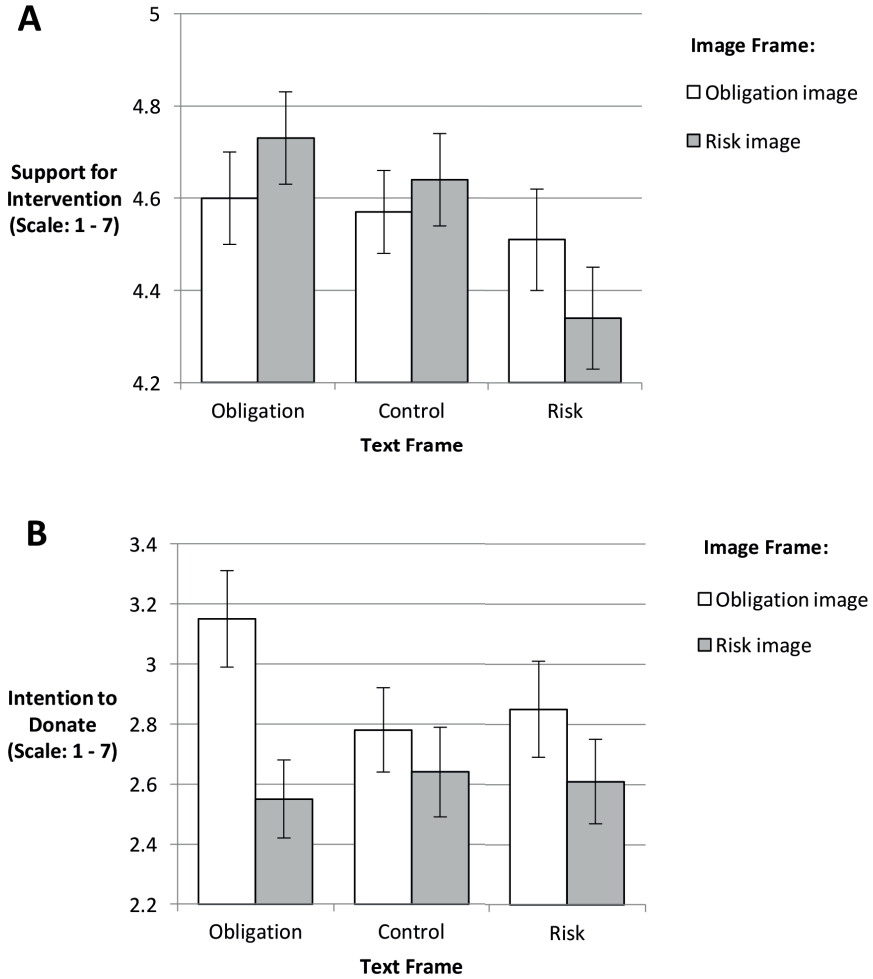


Note. Showing those with relatively low ( $M = 1.19$ ), medium ( $M = 2.78$ ) and high ( $M = 4.37$ ) issue-specific knowledge of the conflict in CAR. An asterisk indicates a significant difference between the text frame conditions,  $p < .05$ .

### Image-text combinations

Results in the image-text combination conditions showed that, counter to our expectations, the frame carried by the text determined participants' support for intervention. Conversely, the frame carried by the image determined participants' behavioural intentions towards the conflict in CAR.

Figure 2.2. Mean differences between the image-text combination conditions for (a) support for intervention and (b) intention to donate.



Note. Panel A shows the main effect of text frame on support for intervention. The difference between the obligation and risk text frames is significant ( $p = .02$ ). Panel B shows the main effect of image frame on participants' intention to donate. The difference between the obligation and risk image frames is significant ( $p < .01$ ). Means and standard errors are plotted. Note that the y-axes on both charts do not reflect the full range of the scales.

There was a significant main effect of text frame on participants' support for intervention,  $F(2, 676) = 3.44, p = .05, \eta_p^2 = .01$ . Post hoc comparisons showed that, across text frame conditions, support for intervention was significantly higher after reading the obligation ( $M = 4.66, SD = 1.03$ ) compared to the risk text ( $M = 4.43, SD = 1.04; MDiff = .23, p = .02$ ). The difference between the control text ( $M = 4.60, SD = 1.04$ ) and the obligation and risk texts were non-significant. In contrast, there was no main effect of image frame on participants' reported support for intervention,  $F(1, 676) = 0.20, p = .89$ , and no interaction between image and text frame,  $F(2, 676) = 1.19, p = .31$ . Therefore, when images are combined with text, participants' support for intervention was influenced by the text frame, with the image frame having little to no effect (see *Figure 2.2A*).

For the behavioural intentions measure, there was trend towards a main effect of image frame,  $F(1,676) = 2.17, p = .14$ . This was significant for the intention to donate item,  $F(1,676) = 7.39, p < .01, \eta_p^2 = .01$ . Participants who viewed a stimulus with an obligation image ( $M = 2.92, SD = 1.65$ ) reported higher intention to donate to the cause in CAR than those who viewed a stimulus with a risk image ( $M = 2.60, SD = 1.45$ ; see *Figure 2.2B*). The main effect of text frame did not approach significance, neither for the combined behavioural intentions scale ( $F(2, 676) = .43, p = .80$ ) nor the individual items. No significant image by text frame interaction was observed,  $F(2,676) = 0.63, p = .54$ .

Taken together, these results partially support hypotheses 2a and 2b, and suggest that: *when images and text are combined, the frame carried by the text influences opinions towards military intervention, regardless of the image frame. In contrast, the image frame determines participants' behavioural intentions, particularly the intention to donate, irrespective of the text frame.*

## Indirect effects of emotions

For images-alone, but not text-alone, mediation analyses showed a significant indirect effect of emotional responses on support for intervention and behavioural intentions.

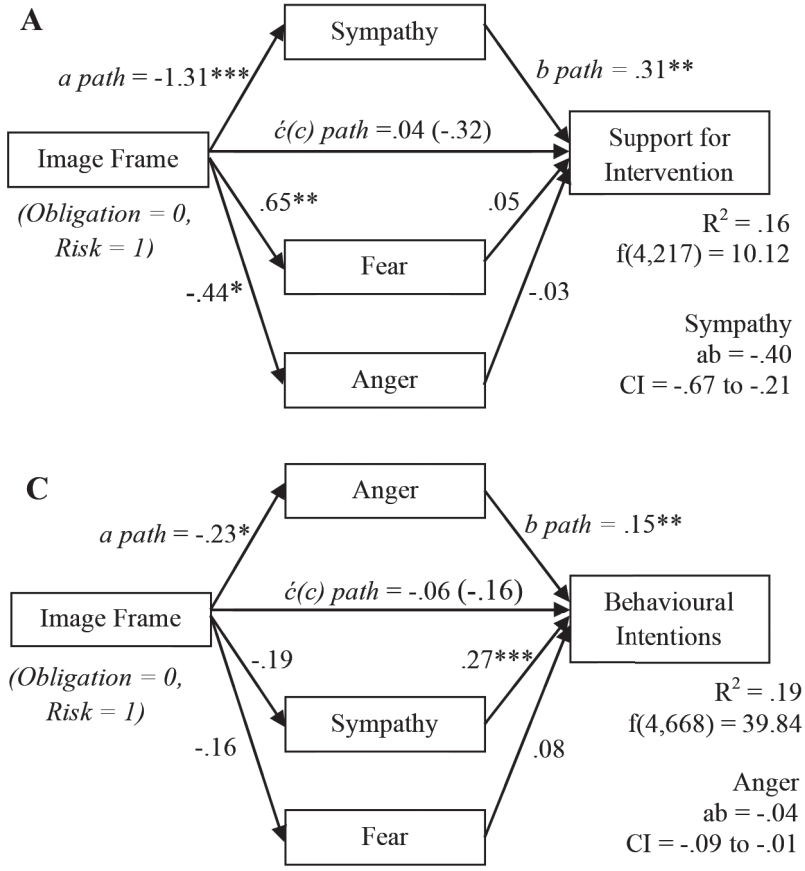
Mediation models for support for intervention and behavioural intentions were tested in the image-alone condition (*Figure 2.3*, panels A and B). A strong indirect effect of sympathy was observed for both measures. Participants who viewed an obligation image reported higher sympathy, and participants who experienced more sympathy displayed higher support for intervention and behavioural intentions towards the conflict in CAR. Furthermore, a significant indirect of fear on behavioural intentions was observed: participants who viewed a risk image reported higher levels of fear, and participants who experienced more fear reported stronger behavioural intentions. When frames were presented via text-alone, there was no indirect effect of participants' emotions on effects for either measure.

For the image-text combinations, mediation analyses showed no indirect effect of emotions for support for intervention. Additional analyses showed that an indirect effect was in fact present but was moderated on the b-path by issue-specific knowledge. Since the purpose of this study was to focus on the previously untested main effects and take a first look how emotions contribute to multimodal framing effects, these results are included in *Appendix A4*.

For the image-text combination conditions, there was an indirect effect of anger on participants' behavioural intentions for the obligation image frame, and an indirect effect of fear on behavioural intentions for the risk text frame (see *Figure 2.3*, panels C and D). Participants who viewed a stimulus with an obligation image reported higher levels of anger, and those who experienced more anger reported stronger behavioural intentions. Participants who viewed a stimulus with a risk text reported higher fear, and those who experienced more fear reported stronger behavioural intentions.

Taken together, these results support hypothesis 3: *indirect effects of emotions on framing effects are stronger for images than text for both the image/text-alone conditions and image-text combinations*. In a similar way to the main effects above, mediation results differ according to the dependent variable. In the combination conditions, indirect effects of emotions are present for behavioural intentions, but are absent, (or, rather, conditional on issue-specific knowledge – see *Appendix A4*) for opinions towards support for military intervention.





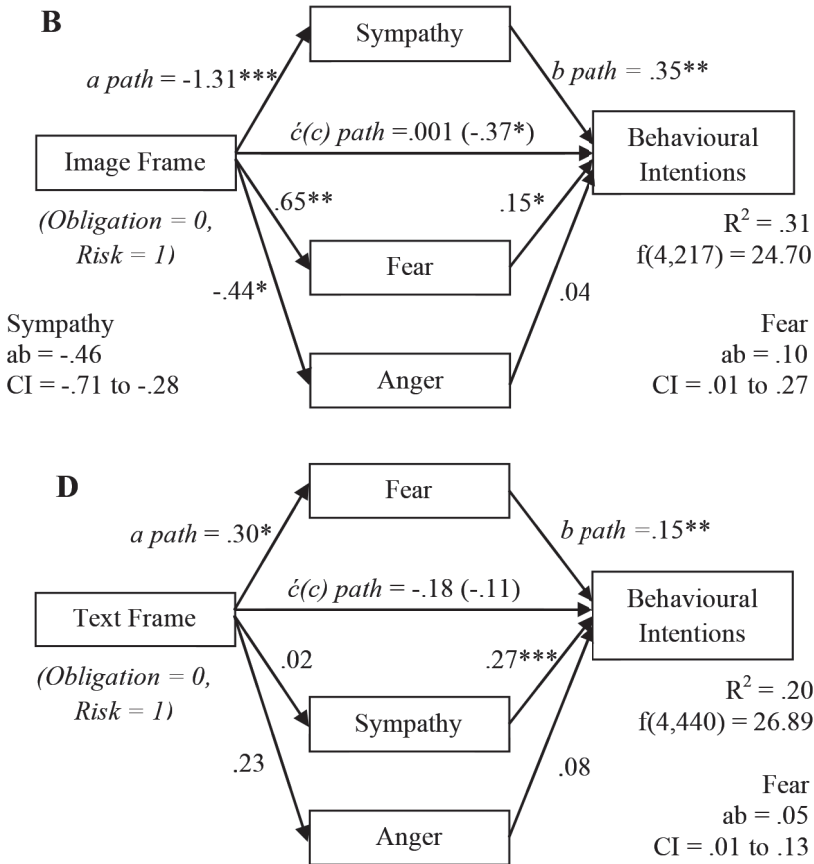


Figure 2.3. Mediation models showing the indirect effects of emotions on framing effects. Unstandardized beta coefficients are shown,  $*p < .05$ ,  $**p < .01$ ,  $***p < .001$ . 95% bias-corrected bootstrap confidence intervals based on 10,000 bootstrap samples are shown for significant indirect effects at the bottom of each panel. Panel A: Indirect effect of the obligation image frame on support for intervention through sympathy in the image-alone condition. Panel B: Indirect effect of the image frame on behavioural intentions through sympathy and fear in the image-alone condition. Panel C: Indirect effect of the obligation image frame on behavioural intentions through anger in the image-text combination conditions. Panel D: Indirect effect of the risk text frame on behavioural intentions through fear in the image-text combination conditions.

## Discussion

This study investigated the role of visuals and text in framing effects using the context of a little known conflict in Africa. The results showed that, when viewed in isolation, images delivered stronger framing effects than text. This was consistent with past research illustrating the power of visuals. In contrast, framing effects for the text-alone condition were only present for those with high issue-specific knowledge.

When images and text were presented in combination, the results suggest a nuanced role of visuals (Domke et al., 2002). The frame provided by the image exerted a dominant influence over participants' behavioural intentions (i.e., their intention to discuss, donate, sign a petition and protest in support of those embroiled in the conflict). In contrast, the persuasive power of text frames determined participants' opinions towards military intervention irrespective of the accompanying image.

Emotions played a leading role in explaining the observed effects. This was particularly the case for images-alone (but not text-alone) where sympathy was powerful in eliciting support for intervention and stronger behavioural intentions. For the image-text combinations, levels of anger and fear predicted framing effects for behavioural intentions, whereas there was no mediating effect of emotions over participants' opinions. Taken together, these findings go beyond previous research (e.g., Iyer et al., 2014) to show that framing effects, and the mediating role of emotions therein, are dependent on the framing device and the dependent variable.

To our knowledge this study is the first to show that images presented alone can have a direct framing effect on policy support. This is an interesting insight because, although seldom viewed in complete isolation, vivid visuals are increasingly prominent in contemporary news formats and have the potential to encapsulate an often apathetic public's enduring perception of an issue (Barry, 1997). The images of victims produced strongest framing effects through sympathy, showing that sharing in the suffering of others predicts helping (Iyer et al., 2014). Therefore, in support of commentators who claimed that powerful media images can affect foreign policy, our results show that images alone can make a small but significant contribution to an interventionist public response.

When framed messages are presented via text alone, prior knowledge is important (Schuck & de Vreese, 2006). The obligation frame failed to resonate in those lacking the cultural knowledge of the conflict and therefore did not trigger support for intervention. On the other hand, this contingency upon knowledge was absent for images-alone, where framing succeeded because the images' key message was clear and participants' emotional reaction was robust (Geise & Baden, 2014). These results are consistent with Sparks et al.'s (1998) proposal, built upon dual-route theories of persuasion (Chaiken, 1980), that images change opinions by more heuristic processing via emotions

compared to text, whereas text exerts a persuasive effect through more systematic, cognitive processing, especially in knowledgeable recipients. The contrasting roles of emotions and knowledge in the visual and verbal framing effects observed require replication, but point towards message modality as a worthy addition to dual-route theories of attitude change.

I now turn to the image-text combination conditions, in which the influence of image and text over framing effects was contingent on the dependent variable. Consistent with our expectations, participants' behavioural intentions towards the conflict in CAR were primarily driven by emotional responses to the image frame. Anger was particularly strong in generating these effects, suggesting that feelings of anger in response to the suffering inflicted on the victims prompted participants to confront a problem by taking action.

In his *Laws of Emotion*, Frijda (1988) emphasised that all emotional states involve some form of "action readiness" (p. 351) which corresponds to impulses, for example to approach or avoid, shout or move. These impulses either remain covert or manifest themselves in overt behaviour, which has been observed experimentally in the context of news media by measuring approach-avoidance responses to TV news images (Newhagen, 1998). I argue that the emotional consequences of visuals in this study elicited a heightened state of action readiness, which in turn resulted in higher behavioural intentions ratings. This contention is also supported by the *affect heuristic* (Slovic et al., 2007), which states that emotions provide a rapid and automatic feeling of the "goodness or badness" of a stimulus (p. 1333), which can guide behaviour. This requires verification, and measuring approach-avoidance behaviours should prove a useful direction in doing so.

In contrast, the influence of text over opinions was contrary to previous research highlighting the powerful influence of visuals. This is a particularly striking result in light of the non-significant effects in the text-alone condition and significant effects in the image-alone condition. Our results suggest that the inclusion of an attention-grabbing image increased the salience of and attention to the accompanying text<sup>7</sup>, whose structure in turn guided participants' interpretation and support for intervention (Zillmann et al., 2009). For opinion formation, it seems that images are important but words "still exert defining power" (Entman, 1993, p. 104).

The absence of a mediating effect of emotions on participants' opinions is also consistent with Sparks et al.'s (1998) assertion that text frames generate persuasive effects through a more systematic, cognitive and less emotionally-driven route than images. Additional analyses in *Appendix A4* lend a more nuanced perspective and further support for this notion. Moreover, Breckler & Wiggins (1989) point out that attitude and opinion change is highly reliant on cognitively elaborated verbal representations. It follows then, that cognitive elaboration of a framed text should be

well-suited to modifying verbally represented opinions, and certainly more so than a framed image. This contention also requires further examination, but is grounded in research emphasising a divergence between cognitive and affective decision making under risky conditions (*risk as feelings* versus *risk as analysis*, Loewenstein et al., 2001). The results also stimulate questions such as how these framing effects might differ if opinions were tested visually instead of verbally, and in those with a highly visual or verbal processing style (Sojka & Giese, 2006).

The use of frames from war and conflict news in this study gave us a wealth of vivid images at our disposal. However, extraneous factors particular to this context warrant consideration. International relations research has shown that public support for military force is determined by a number of dynamics – primarily a war’s objective and its initial success and casualty rates (e.g., Eichenberg, 2005). However, since these factors were kept constant across conditions and, moreover, by matching our stimuli for other potential confounds – including arousal, valence, salience, ambiguity, complexity and credibility – increases the likelihood that our results generalise to other contexts. This, of course, requires further scrutiny and manipulation of constructs such as arousal, valence and ambiguity should provide clues to the prevailing factors in multimodal framing effects.

That said, there are notable limitations of this study. First, the findings concerning the mediating role of emotions can provide only a first estimation of the mechanisms of multimodal framing effects. For instance, the inclusion of participants’ cognitive appraisal of the stimulus (e.g., Lazarus, 1991; Iyer et al, 2014) would have better informed us of the cognitive components of the framing process. Furthermore, I did not explicitly manipulate emotions in our design, and thus strictly cannot make valid inferences about the role of emotions as a causal mechanism of framing effects (Imai, Keele, Tingley, & Yamamoto, 2011). Separately, like any media effects experiment, particularly using a one-shot design, ours sacrifices ecological for internal validity. It is also possible that the experimental context in this study could have had a particular influence over the image-alone conditions for which there was no accompanying text to guide participants’ perception of the stimulus. Although a precise estimate of this is not possible, inspection of the standard deviations in *Table 1* suggest that any bias introduced was probably not systematic, since response variance was comparable in the image-alone and text-alone conditions. Finally, although the observed effects were significant the effect sizes were small and therefore require replication to reinforce our conclusions.

Before concluding, I briefly address the implications of these findings for democratic processes and policy-making that reflects public opinion. Our results chime with previous studies showing that vivid visuals can help typically apathetic citizens to make political decisions through minimally effortful heuristics (e.g., Lau & Redlawsk, 2001). This can be positive when the situation demands an active response, such as charity

appeals for natural disasters, or for raising a grave but little-known conflict up the public agenda (Wanta, 1988). However, deciding complex matters of foreign policy solely on an empathic emotional response is not uniformly desirable. An angry response to the beheading of western journalists, for instance, should not be the single determinant of an intervention against Islamic State. Fortunately, our results suggest that for opinion formation people *are* receptive to nuances outlined in an article's text. Ultimately, normative conclusions about framing effects depend not only on the multimodal characteristics of the message, but also on the audience and media outlet. The interactive effects of these factors is a prime avenue for further research.

To conclude, this study provided striking new insights about how images and text contribute to news framing effects and the mediating effect of emotions therein. The effects observed depended on the framing device and the dependent variable, with the emotional consequences of images playing a formative role. These findings go beyond previous studies and take a first empirical step towards a *multimodal* framing effects theory. In doing so, I provide a clearer picture to those in the newsroom by showing that the effects of powerful visuals on audiences are more nuanced than they appear at first sight.

## Notes

<sup>1</sup> The medium itself also has an impact. Our choice of online news is justified by a Pew Research Centre report (2012) showing that online news sources are now the second most popular in the US (39%), behind TV (55%) and ahead of radio (33%) and newspapers (29%).

ANOVA showed a significant difference in salience ratings between the stimuli types,  $F(2, 1079) = 18.59, p < .001, \eta_p^2 = .03$ . Post hoc comparisons showed that the image-alone condition was judged most salient ( $M = 4.81, SD = 1.59$ ), followed by the image-text combination stimuli ( $M = 4.22, SD = 1.46$ ), followed by the text-alone stimuli ( $M = 3.95, SD = 1.56$ ; all differences were significant,  $p < .05$ ). Therefore participants judged stimuli with an image to be more attention-grabbing.

<sup>2</sup> ANOVA showed a significant difference in salience ratings between the stimuli types,  $F(2, 1079) = 18.59, p < .001, \eta_p^2 = .03$ . Post hoc comparisons showed that the image-alone condition was judged most salient ( $M = 4.81, SD = 1.59$ ), followed by the image-text combination stimuli ( $M = 4.22, SD = 1.46$ ), followed by the text-alone stimuli ( $M = 3.95, SD = 1.56$ ; all differences were significant,  $p < .05$ ). Therefore participants judged stimuli with an image to be more attention-grabbing.

<sup>3</sup> Multivariate ANOVA indicated that randomisation to the experimental conditions was successful. No differences were observed between experimental groups for age, sex, educational background, race, birth country, political orientation, time taken to complete the survey, and instructional manipulation check success rate (all  $p > .09$ ).

<sup>4</sup> For each condition in which the stimulus included an image, the design included a replica condition containing a cropped version of the image. This cropping manipulation was not examined in this study and therefore is not mentioned further. However, since both uncropped and cropped images were newsworthy exemplars of their respective frames, these conditions were included in all analyses. The effects presented in the results section remained in the same direction irrespective of the inclusion or exclusion of the cropped image conditions.

<sup>5</sup> Since the UN has repeatedly called for increased international support to the conflict in CAR, those with an awareness of the conflict may well be inclined to support intervention. Indeed, knowledge of the conflict positively correlated with support for international intervention ( $r = .11, p < .001$ ). Importantly, ANOVA showed that knowledge of the conflict in CAR did not differ across the experimental groups ( $p > .43$ ).

<sup>6</sup> To test whether the difference in magnitude of framing effects between the image-alone and text-alone conditions was significant, we performed a two-way ANOVA for each dependent variable, with stimulus type (image, text) and frame (obligation, risk) entered as between-subjects factors. A significant type-by-frame interaction would reveal a difference in the magnitude of framing effects between stimulus types. For participants' support for intervention this interaction was non-significant,  $F(1, 395) = 0.74, p = .39$ , whereas a significant interaction was observed for behavioural intentions,  $F(1, 395) = 4.18, p = .04, \eta_p^2 = .01$ . This shows that although framing effects were larger for images than text for both DVs, this was only significantly so for participants' behavioural intentions.

<sup>7</sup> In addition to the salience results mentioned in the Introduction section, this assertion is supported by the image-text combination articles being read for, on average, 30s longer ( $M = 122s, SD = 565s$ ) than the text-alone versions ( $M = 92s, SD = 249s$ ). An independent samples t-test of the log-transformed time data showed that this difference approached significance,  $t(841) = -1.33, p = .18$ .