



UvA-DARE (Digital Academic Repository)

Multimodal news framing effects

Powell, T.E.

Publication date

2017

Document Version

Other version

License

Other

[Link to publication](#)

Citation for published version (APA):

Powell, T. E. (2017). *Multimodal news framing effects*. [Thesis, fully internal, Universiteit van Amsterdam].

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, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 4

Video killed the news article? Comparing multimodal framing effects in news videos and articles

This article is under review.

An earlier version of this article was presented at the 67th Annual Conference of the International Communication Association, May 25, 2017 in San Diego, USA.

Abstract

The recent proliferation of vivid online videos captured at the scene of news events begs the question: do news videos have a meaningful impact on citizens' political opinions and behaviours compared to news articles? This was examined in an experiment using carefully matched videos and articles about the European refugee crisis. Counter to expectations, results showed that articles generated stronger intentions to help refugees than videos, and this was mediated by the depth to which the news story was processed. Despite their increasing prominence and intuitively impactful qualities, news videos do not deliver more powerful effects than news articles.

An increasingly visual media landscape provides an ever more vivid insight into news stories. The advent of TV provided a visual connection with news beyond that of newspaper photographs, and the explosion of videos captured via mobile technology mean that journalists more regularly transport audiences closer to the visceral reality of news stories than ever before (Bock, 2016; Kalogeropoulos et al., 2016). In a typical news article, an impactful image alongside a text is known to evoke emotions and *frame* citizens' perceptions of politics (Grabe & Bucy, 2009; Graber, 1996; Scharrer & Blackburn, 2015; Zillmann, Gibson & Sargent, 1999). However, it is not known whether the richer and more psychologically-activating qualities of dynamic audio-visual news videos exert a relatively more powerful influence over political opinions and behaviours. In this study I use an experiment to shed new light on the contribution of visual and verbal modalities to framing effects generated by news videos and articles.

The ongoing refugee crisis in the Middle East and Europe provides a prime example of how journalists, via various news media formats and particularly visuals therein, can frame a political issue. On 2nd September 2015, publication of the photo of the death of Alan Kurdi magnified the plight of refugees and emphasised their role as innocent *victims* in the crisis. In a matter of hours the image went viral and prompted an emotional outpouring of support for refugees, which found expression via a surge in charity donations and petition-signing, and widespread, albeit temporary, policy change across Europe (Vis & Goriunova, 2015; see also Perlmutter, 1998). However, politicization of the issue by the far-right and several violent flash points contributed to the counter-frame depicting refugees as *intruders*. Here, visuals of large numbers of often intimidating and especially male refugees portrayed them as dangerous strangers posing a threat to Europe's safety and cultural and economic achievements. Framing refugees as *victims* (e.g., Van Gorp, 2005) and *intruders* (e.g., Batziou, 2011; Bleiker et al., 2013), both visually and verbally, exemplifies how news media can make certain aspects of an issue more salient (Entman, 1993), which, depending on the reader, serves to change political opinions and behaviours (Scheufele & Tewksbury, 2007).

The media via which these frames were presented throughout the crisis have been diverse. Written articles remain the norm, both online and offline, and are typically accompanied by one or more images. Due to advancements in mobile internet technology and editing software, videos embedded on news websites containing moving images and a verbal narration provide an attractive alternative (Bock, 2016). Although these media formats can often contain substantively the same information, their properties promote a very different reception experience. Intuitively, the dynamic visual flow of videos is more immersive and promotes a richer and potentially more emotionally-gripping depiction of reality than static images that leave storytelling to the conative abilities of the receiver (e.g., Green et al., 2008; Zelizer, 2010). A video's auditory narration externalizes the implied message of the visuals (Messaris & Abrahams,

2001), albeit in a less structured way than the self-paced and more systematic reading of a news article's text (Tukachinsky, Mastro & King, 2011). Despite these seemingly straightforward routines, existing studies comparing static and dynamic media paint a mixed picture about their relative influence over (political) learning and memory, and the underpinning emotional and cognitive mechanisms (e.g., Lang, 1995; Pincus et al., 2016). Importantly for this study, no prior research has systematically examined the contribution of the visual and verbal streams in videos and articles to news framing effects.

To pursue this goal this study used carefully controlled experimental manipulations of a news story about the ongoing European refugee crisis. I presented articles and videos comprising visual and verbal exemplars of victim and intruder frames combined in congruent and incongruent (i.e., matching and non-matching) pairs. Then I measured how participants' information processing depth and emotions influenced their opinions and behavioural intentions to help refugees embroiled in the crisis. In doing so three questions are addressed: Firstly, do news videos generate stronger framing effects than news articles? Secondly, do these different media differentially enhance the impact of the visual and verbal modalities on framing effects? Finally, what role do emotions and depth of information processing play in producing these effects? These insights shed light on the understudied phenomenon of framing effects via news videos, and represent the first systematic study of multimodal framing across different media formats.

Multimodal framing effects

A growing body of research has emphasised the importance of visuals to news media effects (e.g., Arpan et al., 2006; Grabe & Bucy, 2009; Coleman, 2010; Graber, 1990; Pfau et al., 2006; Zillman, Gibson & Sargent, 1999; Wanta, 1998). However, only a handful of studies have focused on the interactive effects of visual and verbal modalities in combination (Coleman & Wu, 2015; Geise & Baden, 2014; Nagel et al., 2012; Powell, et al., 2015; Seo & Dillard, 2016). A common way to examine this is by exploiting the natural variability in the congruence (or "redundancy") – i.e., matching or non-matching – of visual and verbal information in news items (e.g., Coleman & Wu, 2015; Graber, 1996; Fahmy, Bock & Wanta, 2014). Experimental manipulation of visual-verbal congruence allows one to identify the contribution of each modality to media effects. Congruence studies of media learning and memory (e.g., Brosius, Donsback & Birk, 1996; Drew & Grimes, 1987; Reese, 1984; Severin, 1967) have produced mixed findings, with a tendency for improved memory when visual and verbal input are congruent, and intact memory only for the visual stream when incongruent (Lang, 1995).

The few studies that have manipulated visual-verbal congruence to study framing effects show signs of more consistency. Namely, visuals tend to determine behaviours

such as donations and petition-signing (Seo & Dillard, 2016), whilst verbal input influences political opinions (Powell et al., 2015; see also Boomgaarden et al., 2016). Furthermore, studies point towards modality-specific processing pathways, with visuals being processed more emotionally or heuristically, and text more systematically or cognitively (Chaiken & Eagly, 1976; Lang et al., 1999; Sparks et al., 1998; Tukachinsky et al., 2011). Indeed, the drive to position visuals alongside verbal information in political communication follows an increasing recognition of the role of emotions and heuristics in political decision-making (e.g., Lodge & Taber, 2013). However, no studies have examined whether the processing and effects of multimodal news frames might differ in different media, such as article and videos. That is the goal of this study.

Effects of dynamic and static media

Framing effects of news videos and articles can be informed by considering their distinguishing characteristics. Compared to articles, videos are richer, with their moving images providing a closer approximation to everyday visual experience. By indexing reality (e.g., Messaris and Abrahams, 2001) videos are thought to be more immersive and suited to transporting viewers into a story (Green, Brock & Kaufman, 2004; Slater et al., 2006). Moreover, compared to the unbroken narrative flow of a video's audio stream, the self-paced reading of an article's text requires more pro-active and effortful attention-allocation and imagination on the part of the reader.

These unique qualities have mainly been studied in terms of their effects on media learning and memory, which is used to inform this study of framing effects. Prominent scholars in psychology (e.g., Broadbent, 1956) and communication (e.g., Graber, 1990, 1996) and the body of research that followed (e.g., Grabe et al., 2015; Mayer, 2009) point to cumulative effects of visual and verbal streams when presented in parallel as in a typical news video. Dual-coding theory (Paivio, 1971, 1991) states that two sub-systems exist that operate independently – one specialised for processing verbal information and the other for visual stimuli. When both modalities are used, learning should be enhanced. Moreover, since the reception of visuals occurs quickly and holistically (Barry, 1997; Geise & Baden, 2014; Graber, 1996) the addition of visual to verbal input should not burden information processing but benefit learning (see also *Cue-Summation Theory*, Severin, 1967). It follows, then, that attention-grabbing news videos and viewers' ability to process their synchronous visual-verbal streams supports learning and perhaps also produces robust framing effects.

Theories that emphasise information processing requirements over information richness provide an alternative perspective. The Limited Capacity Information Processing Approach (Lang, 1995, 2000) states that visual and verbal streams can facilitate learning as long as the receiver's processing capacity is not overloaded. Lang (1995) synthesised a previously mixed body of research to argue that the structural

characteristics inherent in multimodal TV and video content – including potentially distracting production techniques such as camera cuts and zoom – introduces complexity which can readily overload viewers’ limited processing capacity (Lang, Bolls, Potter & Kawahara, 1999; Lang et al., 1993; see also Wickens, 2002). By contrast, although news articles with a text and image(s) engender serial processing by the viewer, they possess fewer structural constraints and self-paced reception allows the receiver to operate within capacity (e.g., Sundar, 2000; Unnava, Burnkrant & Erevelles, 1994). As such, a limited capacity approach might propose that rich and dynamic news videos contain more unintended obstacles to information processing and learning than articles, and thus will not necessarily exert a more powerful impact on political opinions and behaviours.

To test these proposals, I use only congruent visual-verbal stimuli to compare the relative power of videos and articles in framing effects. Although prior literature on media learning is mixed, the weight of evidence leans towards the intuitively impactful qualities of visceral news videos. Therefore I tentatively propose the following hypothesis:

H1: When visual and verbal streams are congruent, framing effects are stronger in news videos than articles.

Visual and verbal effects in different media

Further inspection of the distinct characteristics of news videos and articles point to potential differences in the relative power of the visual and verbal modality to influence framing effects. I propose that visuals should disproportionately drive effects in videos, whereas text should determine effects in articles.

Videos are primarily visual in nature. A ten second video clip typically contains hundreds of still images which are fused by the visual system into continuous motion. The human brain is finely tuned for perceiving moving stimuli, particularly of other people (Pelphrey & Morris, 2006), in order to provide clues for social interactions and decision-making (van Driel, Grabe, Bas & Kleemans, 2016). Moreover, the vividness of a video’s visual stream mirrors direct experience, something which has been shown to improve media memory, evoke emotional reactions and may be more persuasive (for a review see Taylor & Thompson, 1982). Finally, when learning from congruent and incongruent visual and verbal streams in TV news has been compared, a relatively consistent superiority is observed for the visual stream (Nelson et al., 1976) to the detriment of the verbal modality (Lang, 1995; but see Nagel et al., 2012).

In contrast, news articles may bias verbal input. Although images are attention-grabbing (Garcia & Stark, 1991), the meaning of a news article is primarily deciphered through text (Grimes & Drechsel, 1996) which in turn can influence political opinions

(Powell et al., 2015). Eye-movement routines learned from years of reading experience help readers to automatically decipher the syntactic structure of text (Liversedge & Findlay, 2000). Moreover, articles provide the opportunity for detailed re-reading and resolution of opacities in a manner which is not possible for the audio stream of news videos. Furthermore, static images in news articles tend to possess a degree of connotative ambiguity (Geise & Baden, 2014) which can be clarified by an accompanying text (Müller, Kappas & Olk, 2012). Indeed, compared with their central role in videos, for articles vivid images may serve as a “conceptual peg” on which the more substantive meaning of a news text can be ‘hung’ (Paivio, 1991; Zillmann et al., 2001).

I use a congruence manipulation to unpack the relative contribution of visual and verbal inputs to framing effects in videos and articles. Based on research highlighting the relative bias of videos and articles towards visual and verbal inputs respectively, I propose the following hypotheses:

H2a: Visual framing effects are stronger in news videos than articles.

H2b: Verbal framing effects are stronger in news articles than videos.

Processing routes to medium-specific effects

The differential impact of videos and articles on visual and verbal framing effects begs the question of *how* these effects might occur. Along with dual-coding theory (Paivio, 1991; see also Lang, Potter & Bolls, 1999), neuroscience research has argued for a fast and subconscious route through which our brains initially process visual stimuli (Tamietto & de Gelder, 2010), compared to the necessarily slower and conscious processing of verbal stimuli via neocortex (Pessoa & Adolphs, 2010). This is also theorised in the reasoning and persuasion literature (Chaiken, 1980; Epstein, 1994; Kahneman 2011; Petty & Cacioppo, 1986). In particular in Chaiken’s (1980) *heuristic-systematic theory*, involving persuasion via intuitive versus analytic processing, respectively. Although modality was not a defining aspect of this model, Chaiken and Eagly (1983) proposed that decision cues possessed by visual messages sit naturally within the heuristic pathway, whereas text lends itself to processing that is more systematic.

This processing pathway distinction also conforms to the mode-specific communication potentials of visual and verbal content (Coleman & Wu, 2015; Geise and Baden, 2014), which highlight the idea that visuals are more emotionally processed than the cognitive processing of verbal content (Iyer & Oldmeadow, 2006; Pfau et al., 2000; Sojka & Giese, 2006; Sparks et al., 1998). Pertinent to this study, framing effects research has shown that visual effects on political opinions and behaviours are mediated by participants’ emotional responses (Brader, 2005; Iyer, Webster, Hornsey & Vanman, 2014), whilst verbal effects are moderated by issue-specific knowledge (Powell et al., 2015).

In this study I use the context of the ongoing European refugee crisis to investigate whether the role of distinct visual and verbal information processing pathways are accentuated in different news media. In line with previous studies, I propose that visual frames are well suited to evoke *emotions* which in turn influence political opinions and behaviours. The visual portrayal of refugees as victims should elicit sympathy or guilt which may in turn trigger helping behaviour (Iyer et al., 2014). In contrast, when portrayed as intruders, one may feel fear or disgust towards the presence of unwanted strangers, leading to an avoidance response (Lazarus, 1991). Both victim and intruder visual frames might invoke anger – either from the perceived unnecessary suffering of refugees or the unjust intrusions of foreigners – which should inspire steps to resolve the situation (Kühne, 2014). In this way, specific “emotivational” goals expressed via an individual’s emotional response to news visuals can influence political opinions and behavioural intentions (Frijda, 1988)

In contrast, I operationalise more systematic and cognitive processing of verbal frames as increased *information processing depth* (Craik & Lockhart, 1972; Lang, 1995; Wolski & Nabi, 2000). Deeper information processing has been used to index systematic processing of media messages (Griffin, Neuwirth, Giese & Dunwoody, 2002) and is particularly necessary in processing words compared to visuals (Smith & Magee, 1980). As such, verbal presentation of victim and intruder frames should require comparatively deeper and more elaborative processing than visual input, which in turn might influence political opinions and behaviours. Crucially, the manipulation of medium allows us to test whether the effects of discrete visual and verbal processing pathways are more accentuated in news videos and articles, respectively.

Based on expectations of increased visual and verbal effects in news videos and articles, and integrating this with insights about distinct heuristic and systematic processing pathways for visual and verbal media, I propose the final set of hypotheses:

H3a: Effects of visual frames in news videos are mediated by emotions.

H3b: Effects of verbal frames in news articles are mediated by information processing depth.

Method

Design

An online experiment in the context of the ongoing European refugee crisis was conducted to test the hypotheses. Participants were randomly assigned to twelve conditions in a two (*medium*: article, video) by two (*visual frame*: victim, intruder) by three (*verbal frame*: victim, control, intruder) factorial design.¹

Participants

1123 Dutch adults aged 18 to 75 were recruited via an online data collection panel, Survey Sampling International, in early August 2016. The Dutch case was chosen since the issue is relevant, comparatively so with the rest of Western Europe. Of the original sample, 169 participants dropped out before completing half of the survey or failed an attention-check question. 27 were removed for 'straight-lining' behaviour and 7 were excluded after reporting that the stimulus did not display properly. 7 participants were removed who took unusually short (<5 mins) or long (>3 hours) to complete the survey. The final sample comprised 923 participants, for whom it took a median time of 12 mins 11s to complete the experiment.

The sample was representative of the Dutch population for gender (466 females, 51%) and fairly representative for age ($M = 47.32$, $SD = 16.54$). 93% of participants were born in the Netherlands and 40% had at least one parent who was not born in the Netherlands. There was a range of educational backgrounds and political ideologies.² Participants were reasonably knowledgeable about the refugee crisis (1 = *Not at all*, 7 = *Extremely*, $M = 4.50$, $SD = 1.27$).

Stimuli

Stimuli were selected from media coverage of the European refugee crisis. A rigorous pretesting procedure, explained in *Appendix C1*, helped us to: (1) achieve the victim and intruder frame manipulations in the visual and verbal modalities³; (2) maximise visual-verbal congruence/incongruence; (3) keep the content of the articles and videos the same as far as possible; and, importantly, (4) control for factors known to influence media effects – including perceived arousal, salience, complexity, ambiguity and credibility (Lang, 1995; Severin, 1967; Powell et al., 2015).

Briefly, the following steps were taken in developing the stimuli: Several article texts were downloaded from the *BBC News* and *UN High Commission for Refugee* websites and modified to reflect the victim, intruder and control frames. The victim texts comprised words such as *protect victims*, *suffering*, *help*, *sad*, *sympathy*, *desperate* and *tragedy*. The victim texts included words such as *protect borders*, *burden*, *violence*, *hostile*, *threat* and *flood of migrants*. For the control texts we removed words and phrases

included as part of the frame manipulations. The final framed texts were then used as the script for the audio stream, which were read by a professional broadcast journalist. The videos were made using clips from online news sites (e.g., *The Guardian*, *SkyNews*, *RuptlyTV* and *Human Rights Watch's* YouTube channel). Victim videos included scenes of refugees being pulled from the sea, receiving medical care and sleeping on streets. Intruder visuals included large numbers of refugees near security fences and acting violently towards border guards. Still images were captured from these videos to be used in the articles. The final framed victim image (from *Human Rights Watch*) depicted a young boy recently pulled from a boat, and the selected intruder image (from *Ruptly TV*) showed violent migrants crowding around a fence.

These elements were combined so that the final articles and videos contained the same basic structure: a first section containing basic factual information about development of the crisis, a central part containing the core framed content, and an end section containing factual information concluding by stating the crisis is not yet resolved. This operationalisation is consistent with the definition of news framing as the same factual information presented with a different emphasis (e.g., de Vreese, 2005). In the articles, each of the three sections was each around 35 words in length. For the videos, each section lasted approximately 15 seconds. A still image from all three sections was inserted into the articles so that all parts of the video were visually represented in the articles. The images presenting factual information showed charts of the statistics behind the crisis. The image taken from the first part of the video showed the proportion of the Syrian population affected, and the image from the concluding part showed numbers of asylum applications in Germany and the rest of Europe. These two graphics were added to all versions of the articles and thus were the same for all conditions. Since a typical online news article contains more visual material than merely framed elements, this addition enhanced the external validity of our stimuli whilst ensuring the articles and videos contained the same information as far as possible, thus also maximizing internal validity.⁴

The final selected articles were matched in length (victim text: 112; intruder text: 117; control text: 108), and the videos lasted just under 50 seconds. An example stimulus article can be seen at *Appendix B2* and all stimulus videos can be seen at *Appendix C2*.⁵

Procedure

After entering the survey, participants were asked about their knowledge of the refugee crisis, their preferred political party, and their political orientation.

Participants were then randomly assigned to one of the stimulus videos or articles, which were presented on a blank screen. In the video conditions, participants saw the stimulus video on the page and clicked on it to start watching. Participants were not allowed to progress until at least 55 seconds had elapsed – enough time to view the

video in its entirety. A test question ensured that participants had the sound on before viewing the video. In the article conditions, participants were not allowed to progress before 25 seconds had passed, and the survey automatically progressed after 90 seconds – more than enough time to read the article. Immediately before the stimulus, participants were informed that they would be viewing a news video or article about the refugee crisis and clearly told about the time they were given to do so.⁶

Next, the dependent variables were displayed. Questions measuring support for refugees coming to Europe, behavioural intentions to act in support of refugees, emotions felt when viewing the stimulus, and information processing depth were shown on successive pages. At the end of the experiment participants provided basic personal information, such as age, gender, education level, birth country and parents' birth country.

Measures

After viewing the stimulus, participants answered questions measuring the key dependent variables. Support for the policy of allowing Syrian refugees into Europe was measured using two items: “*There should be more Syrian refugees allowed into Europe*”, and “*Syrian refugees should be prevented from seeking asylum in Europe (reversed)*” (1 = *strongly oppose*, 7 = *strongly support*, $M = 3.82$, $SD = 1.65$, $r = .69$). Behavioural intentions to act in support for Syrian refugees were measured using three items (*intention to: share the news item on social media, donate to charity, sign a petition*; 1 = *very unlikely*, 7 = *very likely*; $M = 2.81$, $SD = 1.38$, $\alpha = .75$).

Questions measuring the mediator variables followed. Participants indicated the extent to which they felt a number of relevant emotions of interest whilst viewing the stimulus (1 = *not at all*, 7 = *extremely*), in the same manner as Iyer et al. (2014). Two items assessed fear (*afraid, anxious*; $r = .58$), anger (*angry, furious*; $r = .73$), sympathy (*sympathetic, compassionate*; $r = .76$), disgust (*disgusted, repulsed*; $r = .64$), and guilt (*guilty, ashamed*; $r = .64$). Processing depth was assessed using four items adapted from Wolski and Nabi's (2000) *Depth of Processing Scale*. These included: *I was motivated to read this article; I paid close attention to each point that was made; I thought deeply about the contents; my mind wandered as I read the article (reversed)* (1 = *strongly disagree*, 7 = *strongly agree*; $M = 4.22$, $SD = 1.15$, $\alpha = .74$).

Analysis

All analyses were conducted on the two main dependent variables – support for refugees and behavioural intentions. Hypothesis 1 was tested using only the congruent visual-verbal frame conditions in a two-way ANOVA – with stimulus frame and medium as between subjects' factors. Hypotheses 2a and 2b were tested using all congruent and incongruent conditions, entered into a 3-way ANOVA, with visual frame, verbal frame and medium as between subjects' factors.

Hypotheses 3a and 3b – indirect effects of emotions and processing depth conditioned by medium – was tested with ordinary least squares path analysis using Hayes PROCESS-macro in SPSS (Hayes, 2013; Model 7). For all moderated mediation models the visual and verbal conditions were tested separately, with the victim, control and intruder frames included as the levels of the independent variable. For emotions as mediators, participants' sympathy, guilt, fear, disgust and anger were included in parallel. For processing depth, the depth of processing scale was entered as the mediator. For each model, stimulus medium was entered as the moderator on the a-path. 95% bias-corrected confidence intervals based on 10,000 bootstrap samples were used for statistical inference of conditional indirect effects.

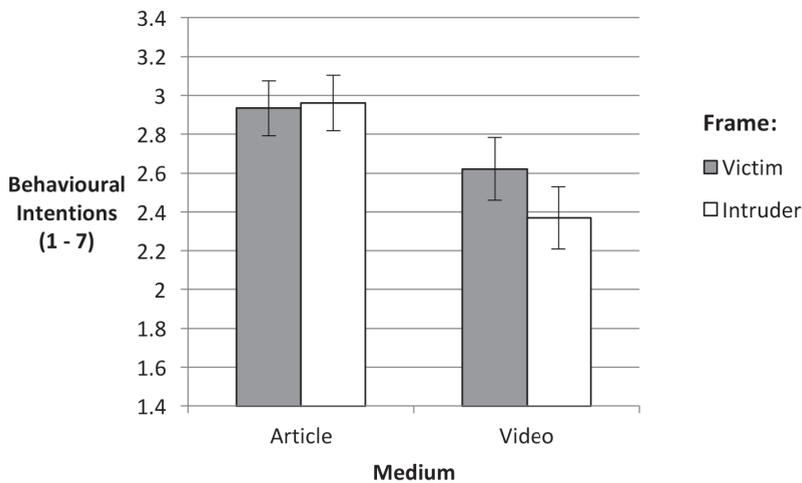
Results

The relative power of videos versus articles

We observed a significant main effect of medium on participants' behavioural intentions to help refugees, $F(1, 302) = 15.47, p = .003, \eta_p^2 = .03$. Participants who read an article ($M = 2.95, SD = 1.29$) reported significantly stronger behavioural intentions compared to those who watched a video ($M = 2.49, SD = 1.34$). No frame-by-medium interaction was observed. See *Figure 4.1*. There were no significant main effects or interactions observed for the support for refugees variable.

These results were counter to the tentative expectations of H1: *Regardless of the frame, those who read an article had stronger intentions to help refugees than those who watched a video.*

Figure 4.1. Mean differences in behavioural intentions between the frame conditions across different media.



Note. Using the congruent visual-verbal frame conditions only. The main effect of medium on behavioural intentions is shown ($p = .003$). Means and standard errors are plotted. Note that the y-axes on both charts do not reflect the full range of the scales.

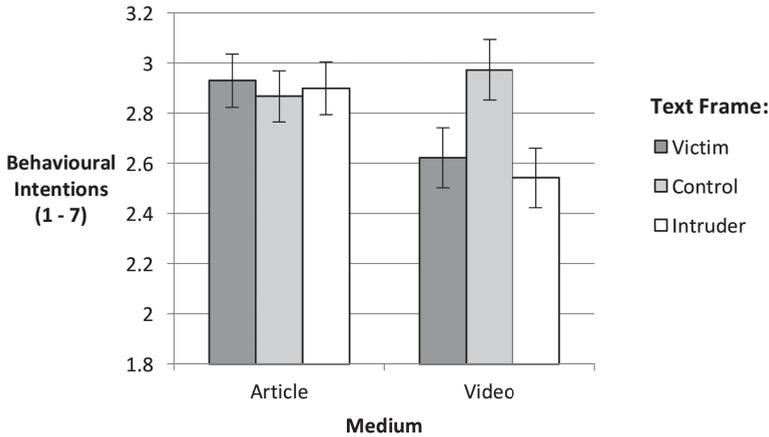
Visual and verbal effects in different media

A main effect of medium was again observed for participants behavioural intentions to act to help refugees, $F(1, 907) = 7.83$, $p = .042$, $\eta_p^2 = .01$. Participants reported higher behavioural intentions to help refugees after reading an article compared to watching a video. This was superseded by a verbal frame-by-medium interaction, $F(2, 907) = 2.56$, $p = .078$, $\eta_p^2 = .01$, albeit just short of conventional significant levels in a two-tailed test, where the effects of the verbal frame was stronger for articles than videos. Stronger behavioural intentions were observed for those who read an article compared those who watched a video for both the victim ($victim_{\text{article}}, M = 2.93, SD = 1.31$; $victim_{\text{video}}, M = 2.62, SD = 1.40$) and intruder ($intruder_{\text{article}}, M = 2.90, SD = 1.32$; $intruder_{\text{video}}, M = 2.54, SD = 1.40$) framed stories. Whereas behavioural intentions were equivalent in those who watched a video or read an article with a control/balanced frame ($control_{\text{article}}, M = 2.87, SD = 1.40$; $control_{\text{video}}, M = 2.97, SD = 1.44$). This is shown in Figure 4.2 and discussed further in the Discussion section. There were no significant main effects or interactions for the support for refugees variable.

These results support H2b: *Verbal framing effects were stronger in those who read an article than those who watched a video.* However, Counter to H2a, no differences were

found in the strength of visual framing effects between those in the video and article conditions.

Figure 4.2. Mean differences between the text frame conditions on behavioural intentions by article and video mediums. The chart shows the main effect of medium ($p = .042$) and the text frame-by-medium interaction ($p = .078$). Means and standard errors are plotted. Note that the y-axis does not reflect the full range of the scale.



Conditional indirect processing pathways

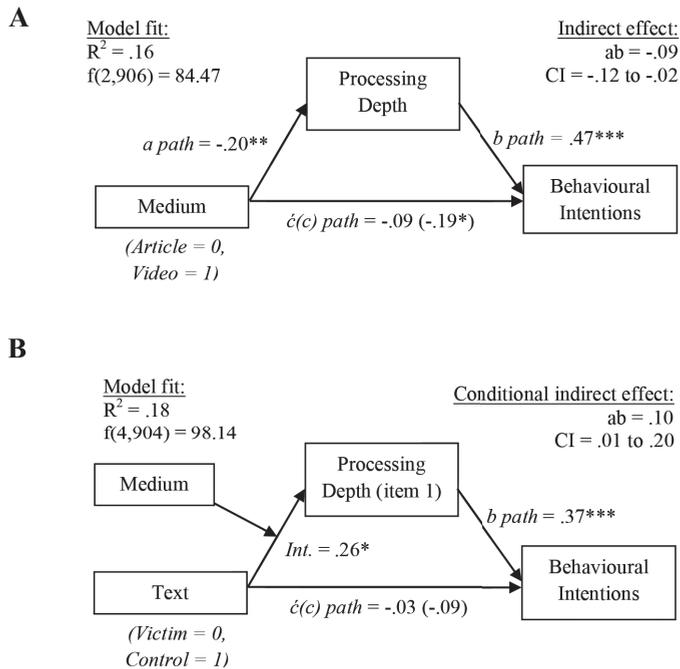
Testing conditional indirect effects to examine H3a and H3b showed non-significant results for the full depth of processing scale and emotion variables. However, a significant conditional indirect effect on behavioural intentions was observed for one of the processing depth measures – “*I was motivated to read/watch the article/video.*” The effect of the victim versus control verbal frame on motivation to process the story was greater in those who viewed an article compared to a video, and this in turn increased behavioural intentions to help refugees (see Figure 4.3B). No conditional indirect effects were found for the support for refugees dependent variable.

In addition, potential mediators of the consistent main effect of medium on behavioural intentions were assessed. To do so, two mediation models were tested with medium (article vs. video) as the IV and processing depth and emotions as possible mediators. A strong indirect effect was observed for depth of processing. Those who read an article processed the story more deeply than those who watched a video, and this increased participants’ intentions to act to help refugees (see Figure 4.3A).

These results provide *partial support for H3b: stronger verbal effects in those who read an article were mediated by their motivation to process the article’s content.* Moreover, the consistent main effect of medium was mediated by processing depth – *those who read*

an article processed the story more deeply than those who watched a video and, in turn, showed increased intentions to help refugees. Counter to H3a, no support was observed for conditional indirect effects of the visual frames via emotions for those who watched a video.

Figure 4.3. Path analysis models showing the (conditional) indirect effects of processing depth on framing effects. Panel A shows the indirect effect of medium on behavioural intentions through processing depth. Panel B shows the indirect effect of text frame on behavioural intentions through processing depth (item 1), conditioned by medium.



Note. 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 top of each panel.

Discussion

I used an experiment employing carefully controlled stimuli about the European refugee crisis to assess the relative impact of different media formats on framing effects. Counter to expectations, those who viewed a news article compared to a video

reported stronger intentions to help refugees by sharing a story, signing a petition and donating money. Compared to the dynamic flow of a news video, the self-paced and pro-active engagement associated with reading a news article (Sundar, 2000; Unnava et al., 1994) led to deeper and more elaborative processing of the story (e.g., Chaiken, 1980; Geise & Baden, 2014), which translated into stronger intentions to help refugees. Moreover, effects of the verbal frame was greater in news articles than videos, and this was also mediated by the depth to which the story was processed (Tukachinsky et al., 2011).

Stronger intentions to help refugees in those who read an article was counter to our expectation that visceral news videos would yield stronger framing effects (Barry, 1997; Graber, 1990; Paivio, 1991; Severin, 1967). Instead, regardless of how it was framed, those who read an article processed the story more deeply, and in turn were more inclined to engage in helping behaviour. This fits with the expectations of limited-capacity models of media effects (e.g., Lang, 1995), which suggest that although videos are information-rich and psychologically-activating, their structural properties such as camera cuts (Lang et al., 1993) that punctuate the parallel stream of visual and verbal input can be distracting and may more readily overload viewer's processing capacity than an article (Lang et al., 1999).⁷ This is in contrast to the self-paced processing of news articles, where attention can be allocated serially and more deeply to the verbal and visual content which, regardless of the frame of the article, delivered stronger effects than the videos in this study. This is interesting for producers of political communication, including parties, campaigns, charities and news publishers who might normally invest resources in producing more vivid but less effective videos.

In addition, results showed that articles accentuated the effects of news frames transmitted via the verbal modality. The syntactic structure of verbal content is particularly powerful in textual form, where well-learned reading routines decode a message's meaning and, unlike videos, allow for complex information to be re-read and disambiguated (Geise & Baden, 2014; Sundar, 2000; Tukachinsky et al., 2011). Going further, our findings suggested that this stronger effect of verbal information in articles is mediated by the depth with which the story was processed (Griffin, Neuwirth, Giese & Dunwoody, 2002). This is consistent with previous studies hinting at the systematic processing of textual content, compared to the heuristic processing of visuals (e.g., Chaiken & Eagly, 1976; Pfau et al., 2000; Powell et al., 2015; Sparks et al., 1998). Compared to dynamic videos, the static depiction of a news story in article form is particularly suited to the systematic elaboration of frames delivered via text, which, in turn, influences behavioural intentions.

A number of findings inconsistent with prior expectations are worthy of further scrutiny. For example, when testing hypothesis 1, we did not observe significant effects

of the victim and intruder frames. This may be due to the refugee issue, which is probably the most salient and politicized news topic of the past two years. Given that factors such as issue knowledge and prior attitudes have been shown to influence framing effects and persuasion (e.g., Arpan et al., 2006; Nelson et al., 1997; Zaller, 1994), a less salient issue might have produced the expected group-level framing effects. Systematic study of medium-specific effects in issues of different salience would help verify this speculation. Separately, vivid news videos did not accentuate framing effects produced by the visual modality. This was unexpected since the moving images of videos are a rich index of real-world visual experiences. Again, it could be that the structural features of news videos distract from the viewer's emotional experience required to accentuate the effect of the visual frames (Lang, 1995), or that scenes from the refugee crisis were too familiar. Alternatively, the uniquely conative quality of still images – i.e., that the events and context that surround a captured image are left to the imagination of the viewer – rival video's moving images in their ability to frame an issue (Barthes, 1978; Zelizer, 2010).

The main limitations of this study are twofold. First, only one political issue was used which limits the generalisability of our conclusions. I chose this approach since the crisis provides a highly interesting context that is well-suited for studying visual and verbal framing. Moreover, importantly, I was able to take utmost care in ensuring that the stimuli were comparable in article and video form, and that the framed stimuli were matched for several factors known to influence framing effects – including perceived arousal, salience, complexity, ambiguity and credibility. An alternative option would have been to use multiple different issues possessing counter-frames which can be readily depicted visually, and then use multiple exemplars of videos and articles from each issue. This approach would indeed improve generalisability, but lacks the control required for this first systematic comparison of multimodal framing effects in different media. A multiple issue and exemplar approach would be an important next step to verify the conclusions of this study. Second, participants were forced participants to view certain articles and videos. In reality, news audiences have the freedom to select content and media formats of their choosing. This may preclude users from selecting articles in the first place (but see Kalogeropoulos et al., 2016) and can influence information processing and effects in ways that cannot be accounted for in this study (Arceneaux, Johnson & Murphy, 2012). Future research should study the selection of different media formats, and compare the effects of videos and articles in forced and selective exposure environments.

This study has several practical implications, particularly for publishers, producers and journalists. Firstly, news articles have a stronger influence over behavioural intentions compared to the same content presented in video form. Therefore political parties, social movements, campaigns or charities who might invest extra resources in creating videos should think again: articles are more likely to influence citizens' political behaviour. Secondly, the verbal content of a framed news item is especially

powerful in news articles compared to videos. Thus if journalists wish to emphasise verbal over visual information then they should use an article. Thirdly, news articles are processed more deeply than news videos. Therefore, if journalists have a complex story that requires an audience's careful consideration then they should deliver it article form, not video.

Finally, this study makes important theoretical contributions. To my knowledge, this is the first study to directly compare framing effects in articles and videos. By drawing on media effects theories from different subfields and through careful control of experimental news items, I take news framing effects into the rarely studied medium of video and, in doing so, set visual and verbal news content on an equal footing. As such, I extend the boundary of *multimodal* framing theory from static image-textual media to dynamic audio-visually. Moreover, when taken together with recent research highlighting online news audiences' preference for articles over video (Kalogeropoulos et al., 2016), these findings reinforce the status of articles, and the inherent flexibility and depth of analysis they provide, as publishers' medium of choice.

To conclude, this study provided surprising new insights into framing effects in different news media and the contribution of visual and verbal modalities therein. Findings showed that news articles were processed more deeply than news videos, and as a result were more likely to encourage political behaviours such as petition signing and charitable donations. These findings provide clear insights for journalism practice and multimodal framing theory by showing that, despite their increasing prominence, news videos do not deliver more powerful effects than news articles.

Notes

- ¹ The experimental groups did not differ on age, gender, education, religious background, place of birth, nor political ideology (all $p > .1$).
- ² 21% were low educated, 37% were middle educated, 42% were high educated. Political ideology was normally distributed around 'moderate' (1 = *left*, 7 = *right*; $M = 3.95$, $SD = 1.52$).
- ³ Success of the frame manipulations was also confirmed in the final experiment using two questions: "to what extent did the video/article portray refugees as 1) victims, 2) intruders?" Two-way ANOVA with two-tailed significance testing showed that the visual frames ($p < .01$) and verbal frames ($p = .093$) differed in the extent to which they depicted refugees as victims and intruders. The control verbal frame achieved a good balance between the victim and intruder frames.
- ⁴ In addition, like any other news articles, our stimulus articles also contained a (framed) headline. To reflect this in the videos, a short 5s segment was inserted at the start, in which the headline was heard and the framed still image was seen.
- ⁵ Translations of stimulus materials and measures from English to Dutch were done using the widely accepted translation-back-translation method, with the authors resolving any discrepancies in the translations.
- ⁶ Inspection of the mean viewing times for videos and articles showed that they were viewed for reasonably similar durations, with longer viewing times for videos ($M = 64.95s$, $SD = 17.69$) than articles ($M = 43.01s$, $SD = 17.26$). This difference is to be expected since the minimum limit for the videos was longer than for the articles.
- ⁷ I checked to ensure that the structural features of the videos in this study were not unusually complex and likely to overwhelm participants' processing capacity. Indeed the videos were rated in the pre-tests as slightly lower than 'moderately complex' (1 = *very simple*, 7 = *very complex*). Moreover, complexity ratings for the different media and modalities were similar: video visuals ($M = 3.33$, $SD = 1.63$); article images ($M = 3.49$, $SD = 1.42$); article texts ($M = 3.66$, $SD = 1.49$).