Overlay ads in humorous online videos: It's a matter of timing

Vermeulen, I.; Droog, E.; Burgers, C.

DOI
10.1080/02650487.2019.1621616

Publication date
2019

Document Version
Final published version

Published in
International Journal of Advertising

License
CC BY-NC-ND

Citation for published version (APA):

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.
Overlay ads in humorous online videos: it’s a matter of timing

Ivar Vermeulen\textsuperscript{a}, Ellen Drooga and Christian Burgers\textsuperscript{a,b}

\textsuperscript{a}Department of Communication Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
\textsuperscript{b}Amsterdam School of Communication Research (ASCoR), University of Amsterdam, The Netherlands

\textbf{ABSTRACT}
Humorous video genres (e.g. comedy/skits, pranks, or humorous vlogs and gaming videos) are among the most popular on online platforms like YouTube. The most common way for brands to advertise in these videos is by means of overlay ads. The current paper investigates the effectiveness of such in-video ads. Based on prior studies on multi-modal information processing, we propose that the timing of showing the brand vs. the humorous build-up may be crucial. In two 4-condition (showing the brand at the joke's start, immediately before, simultaneous to, and immediately after the punch line) between-subjects experiments ($N = 109; N = 394$), participants assessed brand recall, brand-logo recognition, brand attitudes, and purchase intentions after exposure to an overlay ad of a soda brand during a humorous (stand-up) video. Results show the importance of timing: confirming our expectations, showing the brand simultaneous to the punch line reduced recall (Study 1) and logo recognition (Studies 1 and 2). In contrast, showing the brand simultaneous to (Study 1) or immediately before (Study 2) the punch line improved brand attitudes and purchase intentions. Results indicate that the optimal timing of an overlay ad during a humorous video depends on the advertiser's goals: increasing brand awareness or improving brand attitudes.

\textbf{Introduction}
Online advertising is an ever-growing market. While in 2007, Google’s advertising revenues were estimated at 22.9 billion US dollars, in 2017 this number had more than quadrupled to 98.4 billion US dollars (Statista.com 2017). Similarly, advertising revenues of Google’s subsidiary video platform YouTube have increased between 20 and 40% per year between 2015 and 2017 (Emarketer.com 2018). YouTube is the world’s second-most popular social media platform after Facebook, with 1.5 billion monthly
active users – more than Instagram, Twitter, and Pinterest combined (Hutchinson 2017). Humorous videos, including comedy and skit videos, prank and fail videos, and humorous vlog and gaming videos are among the most popular YouTube genres (Mediakix.com 2018). Such humorous videos provide an important setting for firms to advertise their brands and products, and reach a vast audience of potential consumers.

However, prior research suggests that advertising in the context of humorous media content provides specific opportunities and challenges (e.g. Eisend 2009; Gulas and Weinberger 2006). For example, advertising next to humorous content provides opportunities for affective spillover: consumers may associate the positive affect generated by the content with the advertised brand or product (e.g. Gelb and Pickett 1983). At the same time, exposure to humorous content may generate high levels of arousal in consumers, which may hamper their processing of the advertised brand or product (e.g. Sanbonmatsu and Kardes 1988).

In the current paper, we focus specifically on so-called ‘overlay ads’, the most frequently used ad type on most video platforms, characterised by appearing in a peripheral position within the video window while it plays (Cheung and Ho 2017). For these ads, we argue that timing might be a key factor in their effectiveness. Building on the notion that jokes are intended to affect perceivers’ affective states (from opening, to set-up, to punch line, and beyond) – it could be expected that (1) the degree to which affective spillover occurs, and (2) the degree to which users are able to process the ad’s information, may differ depending on when the overlay ad is being displayed. We also argue that the optimal timing of an ad in relation to a joke in terms of affective spillover on the one hand, and in terms of brand awareness on the other, is likely to be different.

Below, we shortly review the literature on the use of humour and advertising, and on the mechanisms posited to make it effective. Next, we discuss the literature on two specific mechanisms in more depth, namely on the use of classical and evaluative conditioning in advertising, and on the relation between jocular humour and information processing. Subsequently, we present two experiments (N=109; N=394) which explore how timing of an overlay ad in relation to that of a humorous video affects users’ brand attitudes, purchase intentions, brand recall, and brand-logo recognition.

**Humour and advertising**

Humour is one of the techniques that has received most attention in the advertising literature (Eisend 2009, 2011; Weinberger and Gulas 1992). In most cases, studies focus on the use of humour within the advertisement itself (Gulas and Weinberger 2006). Overviews demonstrate that, in most cases, humour in advertising impacts affective responses. Specifically, humour increases positive affect and decreases negative affect (e.g. Eisend 2009, 2011, but see Study 1 from Das et al. 2015, for an exception). For cognitive responses, results are more mixed. While some studies found that using humour may enhance ad memory, especially for low-involvement products (e.g. Chung and Zhao 2003), others show that using humour may decrease recognition for the paired stimulus (Kim and Vishak 2008; Strick et al. 2009). An integrative meta-analysis conducted by Eisend (2011) suggests that affective and cognitive routes both
play an important role in the effectiveness of humorous advertising, in that humour boosts positive affect but decreases (critical) central processing.

Yet, in the studies mentioned above, scholars focus on humorous advertisements, in which humour and branded information are fused into one message and the joke is often related to the brand in some way (Eisend 2009; Weinberger and Gulas 1992). In some contemporary, online forms of advertising, this is not always the case. A good example is the advertising genre of overlay ads in video-streaming sites like YouTube. Overlay ads are in-video (sometimes called in-display) ads that appear over part of the video content (Cheung and Ho 2017). In contrast to other video ad formats such as pre-roll or mid-roll ads, overlay ads do not interrupt users’ viewing of the online video; the video continues while the overlay ad is shown in a peripheral (usually bottom, or corner) part of the screen (Cheung and Ho 2017). Overlay ads are currently a popular option for advertisers on video platforms such as YouTube, possibly also because the interruption caused by other ad formats may negatively affect ad and brand attitudes (Goodrich, Schiller, and Galletta 2015).

In case an overlay ad is placed within a humorous video (e.g. during a stand-up routine, or a humorous vlog), the relation between the humorous content of the online video and the brand information in the overlay ad should be differentiated from the relation between humour and brand information in traditional advertising. In traditional advertising, humour and brand are often part of the same advertising message (Gulas and Weinberger 2006), while in an overlay ad, the advertising message is different from and superimposed upon the humorous video. We assume that, when paired with humorous messages, overlay ads can have different effects on recipients. Relatedness is a central construct in the advertising literature that explains when and how pairings of different messages (such as humorous video with overlay ads) can impact consumers.

Traditionally, the advertising literature distinguishes three types of relatedness variables (Gulas and Weinberger 2006). The first type is thematic relatedness, which refers to the way in which the humour is related to the brand (Gulas and Weinberger 2006; Speck 1987). Effects of thematic relatedness are mixed: while some studies (e.g. Lee and Mason 1999) indicate that related humour outperforms unrelated humour, other studies suggest that unrelated humour can improve brand and ad attitudes of low-involvement products (Cline and Kellaris 1999). As placement of overlay ads is typically automated through an algorithm that considers user characteristics (e.g. gender, age, interests) rather than context characteristics (Tucker 2014), it can be assumed that the specific brand of an overlay ad is often (relatively) thematically unrelated to the humorous video.

The second type is intentional relatedness, which refers to the issue whether the humour or the branded content is most dominant within the message (Gulas and Weinberger 2006; Speck 1987). For traditional (print) advertising, messages in which the branded content was more dominant than the humorous elements, performed better than messages in which the humorous content was most dominant (e.g., Spotts, Weinberger and Parsons 1997). Yet, a core characteristic of overlay ads is that they are placed within another message, which implies that the (humorous) content of that message is typically more dominant than the branded content within the overlay ad.
The third type is structural relatedness, which refers to the question where exactly the advertising content is placed in relation to the humorous content (Gulas and Weinberger 2006; Speck 1987). As traditional studies on humour in advertising focus on humorous ads in which the branded content and humour are fused within one message, this topic has traditionally received little attention in the literature (Gulas and Weinberger 2006). Yet, advertising scholars who were primarily interested in context effects have demonstrated the importance of this type of relatedness (e.g. Janssens and De Pelsmacker 2005; Kwon et al. 2019; Stipp 2018). A classic study by De Pelsmacker et al. (2002) for instance placed various types of ads in different types of print and television-media contexts. They found that, in print and TV, positively appreciated media contexts led to a higher attitude towards the ad and, in TV only, also to a higher brand recall.

While such studies demonstrate the importance of context, two important issues should be noted. First, classic context studies make a distinction between media content and advertising, in that the advertising temporarily halts the media content, for instance, by including a commercial break in a (humorous) TV segment, which stops the broadcasting of the segment for an ad message (De Pelsmacker et al. 2002; Janssens and De Pelsmacker 2005). An important characteristic of overlay ads is that they do not stop the original segment, but are broadcasted simultaneously with the segment. Second, most studies operationalise context as being about media content, and for instance contrast a humorous to a non-humorous media context (e.g. Kwon et al. 2019; Stipp 2018). As a result, a context element that has received less attention is the timing of the brand message relative to the media content (Stipp 2018). For overlay ads placed in humorous online content, the timing dimension of structural relatedness – which refers to the issue where exactly the overlay ad is placed within the humorous video – may be crucial to its success. The paradigm of evaluative conditioning provides some further guidelines as to why structural relatedness may be a key issue in these cases.

Evaluative conditioning comprises the hypothesis that a specific stimulus (e.g. the overlay ad) is positively or negatively evaluated after being paired with an affective element of unrelated entertainment content (cf. De Houwer, Thomas and Baeyens 2001; Hofman et al. 2010). For instance, attitudes toward commercial brands or products improved when they were paired with favourable characters (Schemer et al. 2008). Also, product preferences changed for the better or for the worse after the product was – for only a single time – paired with positively or negatively evaluated music (Gorn 1982).

The evaluative-conditioning paradigm posits that when an innocuous stimulus (such as an overlay ad) is paired with a rewarding stimulus, the affective response elicited by the latter may transfer to the former. This effect has been shown to occur fairly regardless of elaboration, and in some cases also regardless of stimulus order – that is, both backward and forward conditioning may induce transfer of affect, although effects generally decrease when the association between both stimuli is weaker, for example when they are presented further apart (De Houwer, Thomas and Baeyens 2001).

Several other studies argued, and observed, that forward conditioning, where the innocuous stimulus precedes the rewarding stimulus, is likelier to generate strong effects (e.g. Stuart, Shimp, and Engle 1987). For example, in the long replication history of the aforementioned musical conditioning study by Gorn (1982; see Vermeulen...
et al. 2014 for an overview), authors were often openly skeptical about the effect sizes that Gorn observed using only the weak procedure of one-shot evaluative/simultaneous conditioning, suggesting that such effects would be more likely if a classical forward conditioning procedure is used (e.g. Allen and Madden, 1985).

For humorous videos, which element constitutes the rewarding stimulus may depend on the type of humour used (e.g. jocular humour, slapstick, funny voices, see Buijzen and Valkenburg 2004, for an overview of humour types in media content). One of the most often used types of humour is verbal jocular humour, in which a comedian tells a specific joke. Popular types of YouTube videos that include this type of humour are (parts of) stand-up routines of comedians like Jerry Seinfeld, Michael McIntyre, and Dave Chapelle, and segments from late-night comedy programs like The Daily Show with Trevor Noah, Late Night with Seth Meyers or Full Frontal with Samantha Bee. In such jocular videos, the rewarding stimulus in terms of simultaneous/evaluative or forward/classical conditioning is the punch line of the joke. This thus entails that, based on any of the two conditioning approaches, the timing of the placement of the overlay ad relative to the punch line of the joke could be crucial in establishing a successful transfer of positive affect from the joke onto the brand. Because competing views exist on which timing of the overlay (immediately before, or simultaneous to) the punch line would be optimal in terms of transfer of affect, we will explore the question:

RQ: When placed in humorous videos, what are the effects of timing of overlay ads relative to a joke’s punch line, on (a) brand attitudes, and (b) purchase intentions?

Humour and brand memory

According to Eisend’s (2011) integrative review, humour reduces cognitive processing. Similarly, a recent meta-analysis by Kwon et al. (2019) demonstrates that brand recall is lower for ads placed within a humorous (vs. non-humorous) context. The Limited Capacity Model of Motivated Mediated Message Processing (LC4MP; Lang 2000, 2006) posits that the cognitive resources people can allocate to various (parts of) a communication message are limited, which means that if one element (e.g. understanding the joke) requires a relatively high amount of cognitive resources, a recipient has relatively few resources available for other tasks (e.g. processing the overlay ad). Thus, an important reason why humour may decrease recognition of paired stimuli is that humour is cognitively demanding, leaving less remaining cognitive resources to process the paired stimulus (e.g. Strick et al. 2010).

Humour theories posit that cognitive resources attributed to jocular-humour decoding vary over the course of the joke. The General Theory of Verbal Humour (GTVH, Attardo and Raskin 1991; Ruch et al. 1993) argues that jocular humour depends on the use of two semantic scripts that (a) have to overlap in some way and (b) should be opposed to each other. For instance, a joke by British comedian Michael McIntyre which we used in Study 2 refers to the Google Earth app. McIntyre posits that this app theoretically allows users to see every place on earth (thus setting up a script of adventurous exploration), but is mostly used to look up the location people are most familiar with, which is their own home (thus introducing a contrasting script of non-adventurous familiarity). The GTVH posits that understanding most jocular humour involves incongruity resolution: A joke sets up a certain situation and expectation,
which is subsequently violated in the joke’s punch line to humorous effect. The GTVH also suggests that recipients of a joke have to cognitively elaborate and use their background knowledge to resolve the joke.

This general scheme of jocular humour may have important implications for stimuli shown during the course of a joke. Because processing capacity is limited (Lang 2000, 2006), and highest appeal to cognitive resources is made at the time of the punch line (Attardo and Raskin 1991; Ruch et al. 1993), recipients will likely have relatively few cognitive resources available to spend on a paired stimulus (e.g. overlay ad) presented during the punch line. Understanding a joke’s set-up requires less cognitive resources, leaving more resources to spend on the paired stimulus. Similarly, after a punch line is understood, cognitive resources are freed and can be allocated to the paired stimulus. We expect this effect to influence recall and recognition for brands (Lang 2000, 2006; Lee and Faber 2007) used in overlay ads of video clips with jocular humour; memory effects may be weaker for brands in overlay ads shown during the punch line than when the overlay is shown at other parts of the video.

Another reason why showing overlay ads during delivery of the punch line may be detrimental to brand memory, may be in the affective response elicited by the joke. A funny punch line is likely to evoke an aroused physical response (Sakuragi et al. 2002), which in turn may reduce processing of simultaneously presented information (Sanbonmatsu and Kardes 1988). The classical notion that high arousal leads to reduced memory performance (Yerkes and Dodson 1908), was recently also tested in the context of first shooter video games, in which brand logos were placed. For brand logos peripherally placed in the game, results were consistent with the Yerkes-Dodson effect: if players were highly aroused by a game sequence, brand recognition was lower than when players were moderately aroused (Jeong and Biocca 2012). This also supports the idea that memory for brands shown in overlays during the punchline of a humorous video will be relatively poor. Hence, we expect that:

\[ H1. \text{Overlay ads placed during the punch line of a joke lead to (a) lower brand recall and (b) lower brand-logo recognition than overlay ads placed at the beginning, before, or after the punch line of a joke.} \]

To explore our research question and to test Hypothesis 1, we present two experimental studies assessing the effects of different timings of overlays in a humorous video on brand recall, brand-logo recognition, attitudes and purchase intentions. Study 1 was conducted in 2012, when overlay ads in online videos were still a relatively new phenomenon; Study 2, using almost exactly the same design, was conducted in 2018.

**Study 1**

**Method**

**Pretest**

We conducted a pre-test \((N = 10)\) using Galvanic Skin Response (GSR) measurements to assess participants’ arousal levels while watching six video clips \((+/- 15–35 \text{s each})\) of a comedian telling a single joke during a stand-up routine. The goal of the pre-test was to be able to select a video for which participants’ arousal levels would distinctly
peak during the punch line of the joke. Based on visual inspection of the resulting averaged arousal curve (see Figure 1), we selected a video by Dutch comedian Leon van der Zanden (the intended subjects of the main study were speakers of Dutch).

Main study; participants and design
A convenience sample of 109 Dutch participants (47.7% men, 52.3% women; age ranging from 17–61 years, $M_{\text{age}} = 34.0$ years, $SD_{\text{age}} = 11.2$ years; 68.8% educated at college level) was recruited by a MSc student through e-mail to join an online study on ‘experiencing media’. Over a period of eight days, two reminders were sent, until further participation stalled. Fellow students in the MSc student’s program were not included, because of potential demand artefacts.

Participants were randomly assigned to one of the four conditions of a unifactorial between-subjects design (brand displayed at the beginning of the video, immediately before the punch line, simultaneous to the punch line, or immediately after the punch line)\textsuperscript{1}. Dependent variables were brand recall, brand-logo recognition, brand attitude, and purchase intention.

Procedure
By clicking a link in the e-mail, participants were redirected to a questionnaire in Qualtrics. The entire experiment took about 6 min to complete. After the participants agreed with the terms and conditions of the study (informed consent), they were randomly assigned to one of the four manipulated video clips. All participants watched the same short video clip (35 s) of the Dutch comedian telling a joke about an embarrassing childhood experience. The joke contained a semantic script opposition between appropriate and inappropriate behaviour: He started with a situation in which he learned a Dutch proverb from his mother (appropriate), and followed with a situation in which he applied the proverb literally in an inappropriate manner. To isolate the effect of the timing of an overlay ad in relation to the punch line of the joke, we purposely chose to not thematically relate the product in the overlay ad with the joke made in the humorous online entertainment video. All videos showed an overlay of a can of Sierra Mist soda in the upper right corner of the screen for 2 s. We chose
the upper right corner because it was most peripheral to the action shown on screen. The videos differed only in the timing of the ad: at the beginning of the video (0.06–0.08), immediately before the punch line (0.26–0.28), simultaneous to the punch line (0.28–0.30), and immediately after the punch line (0.31–0.33). Participants could only continue to the next page of the survey after the entire video had finished playing.

In the overlay ad, brand name and brand logo, as well as the product being a can of soda, were clearly visible due to the use of a professional – actual – advertisement photograph. Because we did not want prior brand familiarity to influence our results, we chose a brand relatively unknown in the Netherlands as our focal brand. Sierra Mist soda was never for sale in the Netherlands (except in online stores catering specifically to US expatriates). A search using social media monitoring service Coosto revealed that in the first four months of 2012, 54 Dutch tweets mentioned the brand Sierra Mist, against 19,115 tweets that mentioned Coca Cola in the same period (Coosto.com, accessed Nov. 24, 2018).

After watching the video clip, participants were asked about brand recall, brand-logo recognition, brand attitude, and purchase intentions, and a control variable: attitude toward the video. Subsequently, demographic information was recorded (sex, age, and education), after which participants were thanked and debriefed.

**Measures**

**Brand recall**

Brand recall was measured by asking participants if they had seen a product while watching the video clip. Participants who answered affirmative were asked to describe the product they had seen in as many details as possible. Answers were coded from 0 (no recall) to 5 (excellent recall). Physical attributes of the product were assigned one point if mentioned correctly (packaging: can; product: soda; colour: green or blue/green). Mentioning a part of the brand name (‘Mist’) or an incorrect brand name (‘Sierra West’) was also awarded with one point; mentioning the correct brand name was awarded with two points. To establish intercoder reliability for recall, two coders independently applied the codebook to all answers. Both coders were aware of the effects being tested, but blind to the condition participants were assigned to. After coding, Krippendorff’s $\alpha$ was calculated, using the KALPHA macro for SPSS with 10,000 bootstrap samples (Hayes and Krippendorff 2007). Intercoder reliability was very high (Krippendorff’s $\alpha = 0.98$, 95%CI = 0.96, 0.99; $M = 2.13$, $SD = 1.32$; see the online supplementary materials at the Open Science Framework for the data files). Disagreements were resolved after discussion between the two coders.

**Brand-logo recognition**

Brand-logo recognition (yes vs. no) was measured by asking participants to identify the specific can that was shown in the overlay ad from four different Sierra Mist cans (with other versions of the logo). Seventy-nine percentage of the participants identified the correct can of Sierra Mist.
Brand attitude
Attitude towards the brand was measured by asking participants how they would rate the brand ‘Sierra Mist’ on six items (bad/good, unappealing/appealing, negative feeling/positive feeling, unpleasant/pleasant, disapprove/approve and worthless/valuable) on a 7-point semantic differential scale (Bruner and Hensel 1992; Low and Lamb 2000). The six items were averaged and combined into an index (Cronbach’s $\alpha = 0.93$; $M = 3.93$, $SD = 0.82$).²

Purchase intention
Purchase intention was measured by asking the participants to indicate the probability that they would want to try ‘Sierra Mist’ in the near future, using three items (very unlikely/very likely, very improbable/very probable, and very inconceivable/very conceivable) on a 7-point semantic differential scale (based on Chang and Thorson 2004). The three items were averaged and combined into an index (Cronbach’s $\alpha = 0.84$; $M = 3.46$, $SD = 1.13$).

Control variable
Video attitude. Attitude towards the video was measured by asking participants how they would rate the video on six items (not funny/funny, unappealing/appealing, negative feeling/positive feeling, unamusing/amusing, boring/interesting and unpleasant/pleasant) on a 7-point semantic differential scale (based on Chattopadhyay and Nedungadi 1992). The six items were averaged and combined into an index (Cronbach’s $\alpha = 0.95$; $M = 4.67$, $SD = 1.29$). The average score on the ‘not funny/funny’ differential was $M = 4.81$, $SD = 1.72$; that is, participants rated the video’s humour qualities above the midpoint of the scale.

Results
A 4-factor MANOVA showed a significant overall effect of the timing of the ad Wilks’ $\lambda = 0.69$; $F(12, 270) = 3.38$, $p < 0.001$, $\eta^2_p = 0.12$.

Focusing on the individual dependent variables showed that timing did not significantly affect brand attitudes ($F(3, 105) = 1.75$, $p = 0.16$, $\eta^2_p = 0.05$), although a significant simple contrast was found between attitudes resulting from showing the ad during the punch line ($M = 4.12$, $SD = 0.56$) vs. showing the ad near the beginning of the video ($M = 3.65$, $SD = 0.82$; $p = 0.03$, Cohen’s $d = 0.67$). No other significant pairwise differences were found (see Table 1 for all means and standard deviations).

Timing did affect purchase intentions significantly ($F(3, 105) = 8.88$, $p < 0.001$, $\eta^2_p = 0.20$). Ad timing simultaneous ($M = 4.02$, $SD = 1.05$; $p < 0.001$, $d = 1.29$) as well as right after the punch line ($M = 3.87$, $SD = 1.12$; $p < 0.005$, $d = 1.14$) produced significantly higher purchase intentions than exposure at the beginning of the video ($M = 2.49$, $SD = 1.30$). No other significant differences were found.

For brand recall, differences between conditions were also significant ($F(3, 105) = 4.81$, $p < 0.005$, $\eta^2_p = 0.12$) but quite different. Simple contrasts showed that brand recall following timing the ad during the punch line was significantly lower ($M = 1.42$, $SD = 1.39$) than timing the ad at the beginning of the video ($M = 2.25$, $SD = 1.23$;
A very similar pattern was found for brand-logo recognition $(F(3, 105) = 4.28, p < 0.01, \eta^2_p = 0.11)$. Logo recognition following ad timing during punch line was significantly lower $(M = 0.58, SD = 0.50)$ than timing at the beginning of the video $(M = 0.83, SD = 0.38; p < 0.005, d = 0.56)$, right before $(M = 0.92, SD = 0.28; p < 0.005, d = 0.84)$ or right after $(M = 0.86, SD = 0.35; p < 0.01, d = 0.65)$ the punch line.3

Because age correlated slightly (negatively) with both brand attitudes and purchase intentions $(r = –0.29, p < 0.05; r = –0.27, p < 0.05$, respectively), and video liking correlated slightly (positively) to brand attitudes $(r = 0.22, p < 0.05)$, we also conducted analyses with both variables included as a covariate (MANCOVA). In terms of significance of observed (simple) differences, this analysis yielded exactly the same results as the earlier MANOVA. These analyses can be found in the online supplementary materials.

### Discussion

Study 1 provides support for the claim that timing of an ad relative to a joke’s punch line is an important factor in the ad’s effectiveness. Consistent with the limited capacity model of mediated message processing (Lang 2000), and classical models of memory performance under different levels of arousal (Yerkes and Dodson 1908). We indeed found that recall and recognition resulting from ads that were timed at the moment the punch line was delivered, were significantly lower than recall and recognition resulting from ads timed at other stages of the joke. This confirms H1.

Regarding the transfer of affect, the issue addressed in the RQ, the results suggest that the affective response elicited by the punch line of the joke, may be (mis)attributed to the shown brand. In particular, participants who were exposed to the brand during delivery of the punch line were more positive about the brand (in terms of attitudes and purchase intentions) than participants who saw the brand at the beginning of the humorous video.

At first glance, it may seem counter-intuitive that affective responses to the brand were highest in the same condition where brand memory brand was lowest. However, numerous studies have shown variations of effects where positive attitudes are produced through superficial processing of a brand rather than through deep processing (e.g. Ha and McCann 2008), or where distracting participants from forming an elaborated cognitive response prevented them from forming defensive reactions (e.g. Osterhouse and Brock 1970; Walter et al. 2018).
In applied terms, the results suggest that the optimal timing of an overlay ad in a humorous video may depend on the brand’s strategy. If the focus of an ad campaign is on increasing brand awareness, timing the overlay during a video’s punch line is not a good idea; users’ awareness will be hampered by cognitive and affective processes taking place during punch-line delivery. So, for new, or yet unfamiliar brands, it may be a better strategy to time overlay ads before or after the delivery of the punch line, in order to optimise the ads’ memory effects. However, if the focus of a campaign is on creating positive brand attitudes, the current results suggest that timing the overlay during a video’s punch line might be sensible, because most positive affect seems to be transferred at that moment. This may be the case, for example, for brands that the public is already familiar with – timing the overlay ad on the joke’s punchline might help to elevate brand attitudes, which may potentially translate in higher purchase probabilities in the period following exposure to the ad. Advertisers thus will have to choose a particular timing of their ad depending on the particular goals they have set for an ad campaign.

A limitation of the current study is its sample size. A sample size of around 25 per cell (although commonly advised at the time the study was conducted – Simmons, Nelson, and Simons 2011) limits the range of effect sizes it can reliably (80%) detect to effects of $d = 0.80$ which, in Cohen’s (1988) terms, are ‘large’ effects. However, before we conducted the study, we had no good reason to expect effects of such magnitude, so the study is likely to be underpowered.

A second limitation of the current study is that we did not check rigorously whether participants were exposed to the entire video experience. Even though none of the participants, in any of the open questions, remarked on having experienced technical difficulties (e.g. failing audio), it might be that some did, which would render their data problematic.

Both the small sample size and the lack of a check on technical difficulties warrant a second (replication) study. Another reason for wanting to conduct a replication study, is that at the time of running Study 1 (in 2012), overlay banners in YouTube videos were (in the Netherlands) still a relatively novel phenomenon (Van den Broek 2009). It might be that their relative novelty made the ads more obtrusive to our 2012 participants than they would be for 2018 participants. This may be a cause for the relatively strong effects that we found on brand memory. It is a well-known phenomenon that after a period of exposure, users may generate forms of banner blindness (Burke et al. 2005), which leads to less attention given to banners, and therefore quite limited ad effects on brand memory. Therefore, a conceptual replication of the original 2012 results in 2018 would testify to the robustness of the observation that timing of overlay ads within humorous video content matters.

A final reason to complement Study 1 with an additional study, is checking whether effects observed are cross-cultural. Humour perceptions, as well as the use and effects of humour in an advertising context, may differ across cultures (Toncar 2001) and thus it is good practice to test effects of basic principles across different cultural contexts.

To address these limitations, Study 2 replicates Study 1, but with much higher power, including checks for technical problems, six years later, and on a UK sample.
Study 2

Method

Pretest
We pre-tested three video clips of jocular humour by English speaking comedians using GSR measurements ($N = 8$). Based on visual inspection of the averaged arousal curve we selected a 31 s video by British comedian Michael McIntyre as stimulus material for this study, because it elicited the clearest peak in arousal during delivery of the punch line of the three videos tested.

Main study; participants and design
The 394 English speaking participants living in the UK (31.5% men, 67.5% women; age ranging from 18–75 years, $M_{\text{age}} = 35.7$ years, $SD_{\text{age}} = 12.1$ years; 63% college level education) were recruited through Prolific (www.prolific.ac) to join an online study. Initially 404 participants took part in the experiment but 10 participants were excluded from the final sample due to technical difficulties (sound not working; screen freezing, etc.; see below). They were paid £0.64 to participate in this study. Recent studies show that respondents on Prolific show less demand characteristics than those on MTurk, while they are more attentive than those of CrowdFlower (Peer et al. 2017); in general, participants recruited through crowdsourcing platforms tend to be more diverse (Casler et al. 2013) and attentive (Hauser and Schwarz 2016) than (student) samples often used for online experiments.

The design echoed that of Study 1, participants were randomly assigned to one of four between-subject conditions (brand displayed at the beginning, immediately before, simultaneous to, or immediately after punch line); dependent variables were brand recall, brand-logo recognition, brand attitude, and purchase intention.

Procedure
The experiment was run in Qualtrics and took participants about 6 min to complete. After participants agreed to informed consent conditions, a sound check was conducted in order to check whether participants would be able to hear (as well as see) the joke being told in the video. In this check, participants were asked to type a number that was enunciated in a sound file. One participant, who was not able to type the correct number, was later excluded from the analysis.

After the sound check, participants were randomly assigned to one of the four conditions. All participants watched the same clip of Michael McIntyre telling a joke about Google Earth. Like in Study 1, all videos showed an overlay of a can of Sierra Mist – which has never been for sale in the UK – in the upper right corner for 2 s, either at the beginning of the video (0.04–0.06), or immediately before (0.24–0.26), simultaneous to (0.26–0.28), and immediately after the punch line (0.28–0.30). Participants could only continue to the next page of the survey after the entire video had finished playing. After watching the video clip participants were asked about brand recall, brand-logo recognition, brand attitude, and purchase intention, as well as some control variables (video attitude, joke recall, and technical difficulties) and demographic information (sex, age, education, and country of birth). They were also reminded of the possibility to opt out of the study within 24 h by emailing their Prolific ID. At the
end of the experiment participants were thanked for their participation and debriefed about the goal of the study.

**Measures**

**Brand recall**
Brand recall was measured and coded in the same way as in Study 1. Again, two coders independently applied the codebook to all data to establish intercoder reliability, which was again very high (Krippendorff’s $\alpha = 0.992$, 95% CI = 0.988, .995; $M = 2.58$, $SD = 1.80$); see the online supplementary materials for the data files. Disagreements were resolved after discussion between the two coders.

**Brand-logo recognition**
Brand-logo recognition (yes vs. no) was also measured in the same way as in Study 1. 70.3% of the participants identified the correct can of Sierra Mist.

**Brand attitude and purchase intention**
Attitude towards the brand (three 7-point semantic differentials; Cronbach’s $\alpha = 0.92$; $M = 4.42$, $SD = 1.12$) and purchase intention (three 7-point semantic differentials; Cronbach’s $\alpha = 0.96$; $M = 4.49$, $SD = 1.64$) were measured in the same way as in Study 1.

**Control variables**

**Video attitude.** Attitude towards the video was assessed using an updated measure with four (not at all entertaining/very entertaining, not at all funny/very funny, not at all humorous/very humorous and not at all amusing/very amusing) 7-point semantic differentials (Becker and Haller, 2014); Cronbach’s $\alpha = 0.98$; $M = 5.42$, $SD = 1.43$. The average score on the ‘not funny/funny’ differential was $M = 5.38$, $SD = 1.50$; again, above the midpoint of the scale.

**Joke recall.** To check if the (paid) participants had actually attended to the video, they were asked to recount the comedian’s Google Earth joke in their own words (open answer). We decided not to drop around 10 participants who did not give a satisfactory recount. The reason was that their answers were too ambiguous (ranging from embarrassment ‘I cannot tell jokes!’ to irritation ‘Babbling on about Google Earth’, to disinterest ‘Can’t remember’) rendering any decision to either include in or drop a participant from the main analysis overly subjective.

**Technical problems.** To check whether the video was playing correctly we asked the participants if they experienced any technical difficulties while the video clip was playing (open answer). Some participants reported technical problems (such as the video freezing), in total 9 participants that experienced some form of technical problem (out the 404 initial participants) were excluded, as well as the aforementioned participant who failed the sound check. The remaining 394 participants were included in the final analysis (see the online supplementary materials for more information).
Results

Similar to Study 1, the 4-factor MANOVA showed a significant overall effect of the timing of the ad: Wilks’ \( \lambda = 0.94; F(12, 1024) = 1.96, p = 0.025, \eta^2_p = 0.02 \).

Brand attitudes were significantly affected by the timing of the ad \( F(3, 390) = 3.39, p = 0.018, \eta^2_p = 0.025 \), with a significant simple contrast found between attitudes resulting from showing the ad immediately before the punch line \( M = 4.66, SD = 1.23 \) vs. immediately after \( M = 4.17, SD = 1.14; p = 0.01, d = 0.25 \) – no other significant differences were found (see Table 2 for all means and standard deviations).

Purchase intentions were also affected significantly by ad timing \( F(3, 390) = 3.49, p = 0.016, \eta^2_p = 0.026 \), showing a similar pattern as for brand attitudes. Purchase intentions were highest when the ad was displayed immediately before the punch line \( M = 4.89, SD = 1.59 \), contrasting significantly with timing immediately after the punch line \( M = 4.17, SD = 1.66; p < 0.005, d = 0.44 \), as well as timing at the beginning of the video \( M = 4.40, SD = 1.79; p = 0.033, d = 0.29 \).

For brand recall, differences between conditions were not significant \( F(3, 390) = 0.57, p = 0.64, \eta^2_p = 0.004 \), and no simple differences were observed \( (p > 0.24) \).

Brand-logo recognition also did not show differences between conditions overall \( F(3, 390) = 2.08, p = 0.102, \eta^2_p = 0.016 \), but a significant simple contrast was found between the lowest score \( M = 0.63, SD = 0.48 \), observed for placing the ad during the punch line, and the highest score \( M = 0.79, SD = 0.41; p = 0.016, d = 0.36 \), observed for placing the ad at the beginning of the video.

Like in Study 1, age correlated slightly (negatively) with brand attitudes \( r = -0.12, p < 0.05 \) and purchase intentions \( r = -0.19, p < 0.01 \); age also affected brand-logo recognition \( r = -0.15, p < 0.01 \). Education level correlated negatively with brand attitude \( r = -0.12, p < 0.05 \), and purchase intentions \( r = -0.18, p < 0.01 \), and positively with recall \( r = 0.19, p < 0.01 \). Video attitudes correlated both with brand attitudes \( r = 0.20, p < 0.01 \) and purchase intentions \( r = 0.24, p < 0.01 \). We ran a MANCOVA analysis adding these variables as covariates – outcomes differed from the MANOVA only slightly: the main effect of ad timing on brand cognition was now borderline marginal at \( p = 0.099 \). There were no other differences between the outcomes of both tests.

All analyses, including the latter MANCOVA analyses can be found in the online supplementary materials.

Discussion

Similar to Study 1, results of Study 2 suggest that the effectiveness of overlay ads in a humorous video on e.g. YouTube may partly depend on the timing of the overlay. Regarding brand-logo recognition, results were most similar to those of Study 1: when the brand was shown simultaneously to the delivery of the punch line by the comedian in the video, logo recognition was the lowest of all conditions; however, the differences observed were smaller than in Study 1, and only the largest contrast was significant. This partly confirms H1. For brand recall, the effects remained far from being significant. A possible reason for the smaller effects on memory encountered in Study 2 compared to Study 1, could be that at the time the former study was conducted, overlay ads were still a relatively new phenomenon, and therefore elicited attention more easily.
Regarding our research question about the spill-over of affect from the humorous video to the shown brand, effects of Study 2 differed from those of Study 1. Whereas Study 1 supported an evaluative-conditioning explanation of the results (attitudes and purchase intentions were generally higher when the ad was placed near the delivery of the punch line), Study 2 yielded the most positive attitudes and purchase intentions when the ad was timed just before the punch line. These results suggest a classical conditioning effect, where the innocuous stimulus (the brand) is placed just before the rewarding stimulus (the punch line) for maximum effect. Indeed, there are several reasons why classical conditioning might be superior to evaluative conditioning; the asynchronous timing may implicitly suggest a causal effect (Kim et al. 1996), and it may diminish the obtrusiveness of the pairing of stimuli and therefore decrease demand artefacts in participants (Vermeulen and Beukeboom 2016). Congruently, prior studies comparing different conditioning mechanisms found forward conditioning to be more effective than backward conditioning (Stuart et al., 1987) – which seems to be corroborated by the results of Study 2. Still, the lack of consistency in the conditioning effects between Studies 1 and 2, as well as the small effects obtained in the latter, hamper the confidence with which we are able to draw conclusions about which mechanism is most likely to be at work: evaluative conditioning, classical conditioning, or both.

Study 2 improved upon the most pressing limitations of Study 1 (limited power, and lack of a technical check). It also did the job of generalising some of the findings (in particular those on memory effects of ad placement in humorous videos) to another era (six years later) and to another cultural setting (UK instead of the Netherlands). A possible limitation of Study 2 may be that, compared to Study 1, the comedian featured in the video seemed better known to participants in the study. In their written answers, some of the participants called the comedian in the video by name, which was not the case in Study 1. Perhaps, participants’ previously formed attitudes toward the comedian (and perhaps even the familiarity with the skit) may have influenced their responses to some extent.

Conclusion and general discussion

In order to find out whether the timing of an overlay ad (as used by, e.g. video sharing platform YouTube) in an online humorous video affects consumer responses, we conducted two experiments. Regarding the effects of timing on consumer memory, the results of both experiments were quite similar: showing a brand simultaneously to the delivery of the punch line in the video consistently yielded the lowest scores on brand recall and brand-logo recognition, although the effects obtained in the original Dutch

Table 2. Means and standard deviations of dependent variables, Study 2.

<table>
<thead>
<tr>
<th></th>
<th>Brand attitude</th>
<th>Purchase intention</th>
<th>Brand recall</th>
<th>Brand-logo recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Beginning of video (n = 96)</td>
<td>4.48</td>
<td>1.10</td>
<td>4.40</td>
<td>1.79</td>
</tr>
<tr>
<td>Before punch line (n = 100)</td>
<td>4.66</td>
<td>1.23</td>
<td>4.89</td>
<td>1.59</td>
</tr>
<tr>
<td>Simultaneous to punch line (n = 98)</td>
<td>4.49</td>
<td>0.93</td>
<td>4.49</td>
<td>1.42</td>
</tr>
<tr>
<td>After punch line (n = 100)</td>
<td>4.17</td>
<td>1.14</td>
<td>4.17</td>
<td>1.66</td>
</tr>
</tbody>
</table>

Means with different superscripts (in columns) differ significantly (p < 0.05).
study were more pronounced than those in the later UK study. These results confirmed our expectations that showing a brand during a punch line would yield negative effects on brand memory, because the attribution of resources needed to process a joke’s punch line, as well as the arousal caused by the delivery of the punch line, are likely to hamper processing of simultaneously presented information (in this case, the ad) and thus will hamper ad memory. Given these results, it is straightforward to give advice to advertisers who aim to leverage the popularity of humorous online videos on YouTube to increase brand awareness: such advertisers should prevent their brand from being shown simultaneously to the delivery of a joke’s punch line.

Effects on attitudinal variables, brand attitude and purchase intentions, were not entirely consistent between both studies presented in this paper. Whereas in Study 1, scores on both attitudinal variables were most positive when the ad was shown simultaneously to the punch line, in Study 2 they were most positive when the ad was shown just before the punch line. Both findings are congruent with a conditioning explanation, where affect elicited by the video is transferred to the brand shown in the ad, resulting in more positive brand perceptions. However, whereas the findings in the former study are congruent with evaluative conditioning, findings in the second are congruent with classical conditioning. This presents a challenge for both theoretical and practical reasons. Theoretically, the current results make it difficult to decide which of the two mechanisms best explains how affect is transferred from the entertainment media to the brands and products in the accompanying ads. Practically, the current results make it hard to advise advertisers on when to display their brands relative to the humorous build-up in a video.

As usage of video-sharing platforms has increased, and overlay advertisements are one of the most common ad types on these platforms, it seems worthwhile to continue the investigation of their effectiveness. Future researchers could build on the current theoretical framework, which considers the effects of differential experiential outcomes of entertainment media (positive affect, arousal, cognitive load) on differential advertising outcomes (brand recall and brand-logo recognition; brand attitudes and purchase intentions). In our experiments, we purposely used a product that was unrelated to the joke, because this resembles the use of overlay ads in practice. Because humour can also be thematically related to the advertised product (and advertising studies show that related humour can outperform unrelated humour in advertising effectiveness; Lee and Mason 1999), the effects found in the current experiments might be amplified by using a joke and a product that are thematically related. Future research should take this factor into account.

In addition, further research into humour and advertising in complex, multi-modal media environments, could take motivational factors stronger into account. While watching humorous videos, quite likely viewers will be more motivated to process the primary content (the joke) than to process any secondary content (e.g. ads). Studies into the mechanisms underlying banner blindness, indeed showed that participants with a different goal orientation may also differentially attend to banners (Heinz et al. 2013). Perhaps such motivational factors could also affect users’ attention attributed to secondary content during the course of a joke told on video.

While our current study focuses on humour, follow-up studies could also consider the effectiveness of timing of overlay ads in relation to other forms of video content.
These studies could also focus on the role of arousal across video types, to establish whether effects of overlay ads are driven by timing and video type, or by general arousal. In addition, our studies were conducted within two different national cultures (the Netherlands vs. United Kingdom). Current research on the role of culture has demonstrated that humour shares certain cognitive structures across cultures, even though its actual content may differ (Alden et al. 1993). Surprisingly, humour may not necessarily reflect cultural values that are dominant within a specific culture (Laroche et al. 2014), and specific sub-cultures within a larger (national) country may respond differently to various humour types (Kuipers 2006; Rößner, Kämmerer and Eisend 2017). Specifically for overlay ads which are targeted towards individual users, future research could further unpack the impact of specific sub-culture membership on the effectiveness of an overlay ad under various conditions of placement.

Additionally, future research could investigate possible differential user responses depending on factors such as (1) whether the ad promotes low or high involvement products, (2) the informational complexity of the overlay ad, (3) the disruptive character of the overlay ad (e.g. depending on size, colour, or movement) (4) users’ habituation to overlay ads appearing, and (5) differences in fit between the advertised product and the humour content. All these factors could affect both memory and affective outcomes of overlay ads appearing during humour sequences.

Although we developed our theoretical framework to apply to advertising within humorous entertainment, it might also apply to other types of entertainment: genres like music, drama, suspense, horror, and even news, are also characterised by intros, set-ups, build-ups, and climaxes. Very likely it holds for such genres as well that advertisers may optimise the effectiveness of their ads by carefully choosing the timing of their ad. Even though such effects might be relatively small (as suggested by the second, better powered, study in this paper), given the sizes of the audiences that many videos have these days, they still may have considerable impact on brand awareness, brand attitudes, and eventually on sales.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Notes**

1. We originally included a fifth condition, placing an ad in the middle of an unrelated clownish video, specifically selected for not containing a punch line. We intended to use it as a control, but already during data collection realised it was unsuitable for that purpose, as it differed from the other videos in many aspects – length, subject matter, entertainment value, etc. Hence, we ignored the data from this condition (22 ppts).
2. We also included an open question on brand attitudes, but this measure turned out to in part also measure brand recall (“I did not see a brand”) rendering it an invalid measure of its intended focal concept; hence, we did not analyse this item.
3. We realise that, due to its dichotomous nature, the variable brand logo recognition strictly does not comply to requirements for testing using a MANOVA. However, joint testing of dependent variables in a MANOVA has advantages in terms of conciseness and interpretability (where, for dichotomous variables, average scores should be interpreted as
proportions). Separate analysis of brand logo recognition using Chi-square tests (for both Studies 1 and 2) yielded identical results, and can be found in the online supplementary materials at the OSF.

4. A possibly interesting observation is that mean scores for brand attitude and PI were higher in the UK than in the Dutch sample. Possibly, transfer of affect from video to brand was more successful in the UK sample. Attitudes toward the humour video were also more positive in the UK sample (5.42 vs. 4.67; \( p < 0.001 \)) and possibly this more positive affective response translated to more positive brand attitudes. However, the difference may also be explained by sample, cultural, and language differences. Perhaps the UK sample had a more positive base-rate attitude toward soft drinks, or a tendency to answer questions slightly more positively. Importantly, because in both samples average scores were still in the center of the score range, it is unlikely that this observed difference influenced results by, e.g. causing floor or ceiling effects.

**Funding**

This study was supported by the Netherlands Organisation for Scientific Research (NWO/VIDI grant 276-45-005).

**Notes on contributors**

*Ivar Vermeulen* is an Associate Professor in Communication Science at the Vrije Universiteit Amsterdam, specializing in persuasion research. Vermeulen obtained his PhD in 2003 at the Institute for Logic, Language, and Information (University of Amsterdam). Most of his research is experimental in nature and focuses on the role of trust, biases, and affective cues (such as music, or humor) in persuasion processes. Other projects involve applications of computational research methods such as automated content analysis, network analysis of Web data, and large-scale meta-analysis.

*Ellen Droog* is a PhD Candidate in the Department of Communication Science at Vrije Universiteit Amsterdam. Her PhD project is funded through NWO VIDI project 276-45-005 and focuses on the use and cognitive processing of figurative language and figurative images in satiric news.

*Christian Burgers* is an Associate Professor in Communication Science at Vrije Universiteit Amsterdam and a Full Professor by special appointment in Strategic Communication (Logeion Chair) in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. In 2010, he defended his PhD dissertation in the field of Communication and Information Studies at Radboud University Nijmegen. His research focuses on the role of language in strategic communication, with a particular focus on figurative language such as metaphor, hyperbole and irony. He also has an avid interest in the use of humor in a persuasive context.

**References**


Yerkes, R.M., and J.D. Dodson. 1908. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology* 18, no. 5: 459–82.