Published in final edited form as:

Int J Advert. 2017 ; 36(1): 82–106. doi:10.1080/02650487.2016.1196030.

Eye-Tracking Evidence that Happy Faces Impair Verbal Message Comprehension: The Case of Health Warnings in Direct-to-Consumer Pharmaceutical Television Commercials

Cristel Antonia Russell, PhD^{1,*}, John L. Swasy, PhD¹, Dale Wesley Russell, PhD², and Larry Engel, MFA³

¹Kogod School of Business, American University, 4400 Massachusetts Ave, NW, Washington, DC 20016, USA

²Uniformed Services University of the Health Sciences, F. Edward Hebert School of Medicine, 4301 Jones Bridge Road, Bethesda MD 20814, USA

³School of Communication, American University, 4400 Massachusetts Ave, NW, Washington, DC 20016, USA

Abstract

Risk warning or disclosure information in advertising is only effective in correcting consumers' judgments if enough cognitive capacity is available to process that information. Hence, comprehension of verbal warnings in TV commercials may suffer if accompanied by positive visual elements. This research addresses this concern about cross-modality interference in the context of direct-to-consumer (DTC) pharmaceutical commercials in the United States by experimentally testing whether positive facial expressions reduce consumers' understanding of the mandated health warning. A content analysis of a sample of DTC commercials reveals that positive facial expressions are more prevalent during the verbal warning act of the commercials than during the other acts. An eye-tracking experiment conducted with specially produced DTC commercials, which vary the valence of characters' facial expressions during the health warning, provides evidence that happy faces reduce objective comprehension of the warning.

Keywords

Direct-to-consumer advertising; eye-tracking; processing; understanding; visual; auditory; congruency; modality

Introduction

Extant cognitive psychology and advertising research has shown that the various components of audiovisual stimuli compete for viewer attention and that one component can distract from another component, resulting in a learning deficit (Kahneman 1973; Waldmann and Holyoak 1992). Television (TV) commercials simultaneously communicate information

*Corresponding author: 1-202-885-1929; russell@american.edu.

in the two competing modalities. Because synchronized audiovisual events capture attention in an automatic manner (Matusz and Eimer 2011), when the audio and visual information match, this temporal synchrony improves understanding of the main message. However comprehension of the information presented may instead suffer when positive laden visual elements accompany a negative audio message, such as a risk or disclosure statement (Sengupta and Fitzsimons 2004).

Concern about cross-modality interference is particularly pertinent in the area of warnings, disclaimers, and disclosures that are designed to explain or qualify an advertising claim and correct inaccurate comprehension, which are often based on inferences from visual imagery. In light of the research findings surrounding the processing of structural elements in advertising, the European Union instituted the Unfair Commercial Practices Directive (Castendyk and Dommering 2008) while the United States (US) developed a clear and conspicuous standard (Hoy and Andrew 2004): both aim to efficaciously present disclosures in advertisements and strengthen disclosure remedies in deception and unfair advertising cases. Although such efforts help protect consumers, research has shown that written warnings or disclosure information is only effective in correcting consumers' judgments if enough cognitive capacity is available to process that information (Johar and Simmons 2000). Further, given that disclosures and warnings often compete with other elements of the advert presented in another modality, they are often unproductive or may even be counterproductive (Amyx, Bristow and Robb 2004; Morgan and Stoltman 2002).

This research is situated in the context of Direct-to-Consumer (DTC) pharmaceutical advertising in the US, where government regulations require that DTC advertisements present a drug's benefits and risks in an unbiased and truthful manner. There are concerns that consumer understanding of the health benefits and warnings presented in DTC advertisements is problematic (Food and Drug Administration 2010; Hollon 2005; New Zealand Ministry of Health 2011; Wolfe 2002), especially for verbal messaging that might be hindered by the many positive visual elements conveyed at the same time as the audio risk information. Such concerns have been expressed globally and underlie policies that prohibit DTCs advertisements; the US and New Zealand remain the only countries that sanction DTCs. Cross-modal interference is thus ripe for testing in the DTC advertising context by experimentally testing whether positive character facial expressions hinder comprehension of the verbal risk warning presented simultaneously. More generally, this research offers a strong empirical test of how the interplay between visual and auditory information in TV commercials affects cognitive processing and message comprehension.

Literature Review

Cross-Modality processing

Building on Paivio's (1971) seminal research on coding redundancy, multiple studies have demonstrated that communicating the same information across two modalities of presentation improves people's ability to process it. Humans have a natural preference for cross-modal correspondence, both from a perceptual standpoint (e.g., matching high pitch sound to bright colors; Marks 1987) and from a semantic standpoint (e.g., matching a pleasant visual with pleasant accompanying text; Spence 2011). Stimuli presented in

different modalities, but semantically or perceptually congruent, tend to be bound together in information processing in what has been referred to as the *unity effect* (Vatakis and Spence 2008). Auditory messages can reinforce a matching visual, as in the case of product placements presented both visually and auditorially (Russell 2002). Written information can reinforce auditory information; for instance, in the context of TV commercials, text subtitles that visually reiterate the information conveyed in the auditory channel attract the viewers' attention and enhance recall and memory of the auditory message (Brasel and Gips 2014).

Surprisingly little research has addressed the potential interference between valence-incongruent messaging (i.e., positive visual images and negative audio messaging). Yet, audiovisual incongruence is frequent; for example, broadcast news coverage often relies on standard or indirectly related visuals to accompany the verbal information, leading to low levels of text-visual correspondence (van der Molen 2001). Although verbal messages often accompany visual information in the advertising context, research on cross-modality interference in the context of advertising is scant (Hertrich, Dietrich, and Ackermann 2011). An exception is a study that documented how the presence of music or visuals distracts from learning and remembering ad copy resultant of the interference each mode faces in perceptual memory's acoustic and visual stores (Tavassoli and Lee 2003). Research on cross-modality processing is growing at a much faster pace in cognitive psychology, especially within neuroscience (Donohue et al. 2013).

Perceptions of facial expressions

Advertising practitioners and researchers alike have long recognized the importance of visual rhetoric in advertising (Scott 1994; McQuarrie and Phillips 2008). Visual executional elements can trigger not only cognitive responses but also affective responses that can mediate the acceptance and efficacy of an advertisement (Batra and Ray 1986). Generally, content analyses of media images, and advertisements in particular, have considered characters' broad demographic characteristics (e.g., gender, race, or age) but few have codified and analyzed facial expressions, despite the known impact of facial depictions on viewers' perceptions (Kolbe and Albanese 1996). One exception is an analysis of articles concerning breast cancer, which found primarily positively valenced faces accompanying the texts (McWirth, Hoffman-goetz, and Clarke 2012).

A large body of psychology research has demonstrated that facial expressions universally convey certain emotions. Facial cues readily transmit prototypical emotions such as happiness, sadness, fear, or anger (Hareli, Shomrat, and Hess 2009), which trigger systematic inferences and social judgments in the receiver (Etkoff and Magee 1992). For instance, one study found that sad faces on charity advertisements led to more donations in response to the plea in the print advert because the readers felt sadder at the sight of the sad face and were more sympathetic to the cause (Small and Verrochi 2009). Similarly, another study demonstrated that consumers automatically make valenced judgments of composite faces generated from morphed images of likeable or unlikable celebrities, even though they are not aware of the origin of these facial cues (Tanner and Maeng 2012).

An empirical evaluation of consumers' responses to facial cues in print advertising found that features of faces affect several metrics of interest to advertising practitioners: attitude

toward the advert, attitude toward the brand, and even intentions to purchase the advertised brand (Xiao and Ding 2014). Faces that express trustworthiness, competence or and attractiveness are especially impactful. The influence of facial expressions is not limited to human facial expressions as recent research has found that an avatar's (i.e., a digital human representation) facial expression can also influence causal attributions and in turn impact consumer behavior (Lee, Kim, and Peng 2013). But, positively valenced visuals such as happy facial expressions may interfere with information communicated simultaneously via the auditory channel, and generate potential issues of miscomprehension or at least, hindered comprehension of any negatively valenced verbal information communicated.

Interference between positive and negative information

Humans have a finite cognitive capacity and therefore selectively attend to and process environmental stimuli (Kahneman 1973). Cross-modality stimulation is especially high in audiovisual media, where visual and auditory cues compete for attention. In this environment, congruence between different sensory inputs facilitates cross-modal processing whereas incongruence between visual and verbal information renders processing more difficult (Hertrich, Dietrich, and Ackermann 2011). Competing audio and/or visual cues can counter each other, thus resulting in a learning deficit (Donohue et al. 2013; Waldmann and Holvoak 1992). Speech understanding is negatively affected by listening effort, which itself increases as a function of the general load on one's attention capacity from all competing stimuli (Kahneman 1973).

The integration, judgment, and understanding of messages are also affected by whether verbal and visual emotion is conveyed in a congruent manner. For instance, simultaneously presenting a face communicating mismatched emotional expressions to the accompanying auditory message of a semantically neutral speech utterance (e.g., a happy face with a sad tone) is sufficient to bias perceptions of the emotional content of the speech: a happy face reduces judgments of how sad the verbal message is compared to a sad face (De Gelder and Vroomen 2000; Massaro and Egan 1996). Clearly, the extant cognitive psychology literature indicates that incongruence in the valence of visual and auditory information renders processing of the message more difficult and could impair understanding of a verbal warning message.

Measuring cognitive processing load

New technology affords researchers with non-invasive physiological measures of cognitive effort during speech or visual perception (Sirois and Brisson 2014). In particular, pupillometry provides a reliable measure of responses to visual stimuli and it is generally accepted that relative changes in pupil size reflect systematic changes in cognitive exertion (De Gelder and Vroomen 2000; Massaro and Egan 2006). Pupil size, which reflects activity in the sub-cortical locus coeruleus attention system, fluctuates as an adaptive response to both changes in the environment as well as changes in cognitive task load (Laeng, Sirois, and Gredebäck 2012). There is evidence that pupils dilate in difficult processing environments, such as when translating difficult words or processing syntactic ambiguities (Engelhardt, Ferreira, and Patsenko 2010; Hyönä, Tommola, and Alaja 1995; Sirois and Brisson 2014).

Pupillometric methods are useful in cross-modal environments because pupil size increases as listening task demands increase (Koelewijn et al. 2012; Zekveld et al. 2014; Zekveld and Kramer 2014). Pupillary response is sensitive to the congruency between stimuli as shown in a classic Stroop color-naming task (Laeng et al. 2011; Laeng, Sirois, and Gredebäck 2012). Incongruence between visual and auditory information may thus be detectable in terms of cognitive load, as reflected by pupil dilation. Indeed, research suggests that pupil dilation can serve as an indicator of listening effort and difficulty of processing speech (Kuchinsky et al. 2013, 2014). In the context of TV commercials, pupil dilation can thus serve as a signal of listening difficulty (Kuchinsky et al. 2013, 2014).

Conceptual Model

The conceptual model, presented in Figure 1 connects the above reviewed bodies of research on cross-modality information processing and its impact on understanding of a message. Based on prior research on cross-modality interference, the model posits that the co-presence of positive visuals and negatively laden auditory information in TV commercials creates a cross-modality incongruence situation that increases cognitive load when exposed to the message (as measured by pupil dilation) and ultimately reduces consumers' ability to understand the auditory message, resulting in lesser comprehension. This model is tested experimentally in the context of pharmaceutical drug advertising.

Research Context

The issue of cross-modal interference is especially pertinent in DTC pharmaceutical advertising. Drug companies spent approximately \$4.5 billion in the US market in 2014 (Kantar Media 2014; Mack 2014). DTC TV commercials follow a common to advertising format: informational content, including the government mandated health warning, is generally presented orally and affective content is conveyed visually, through the characters' facial and body gestures and other visual aesthetics. Existing content analyses reveal that approximately 95% of DTC TV advertisements are premised on emotional appeals (Frosch et al. 2007). While sponsors of prescription drug advertisements are not prohibited from using emotional elements, the US Food and Drug Administration's (FDA) Office of Prescription Drug Promotion requires that DTC advertisements present a drug's benefits and risks in an unbiased and truthful manner. Many have expressed concern that consumer understanding of the health warning, communicated verbally, is hindered by the many positive visual elements conveyed at the same time as the audio risk information (Donohue, Cevalco and Rosenthal 2007). Yet the extant literature does not answer whether and to what degree positive visual elements in DTC advertisements impair consumers' understanding the verbal risk statement.

Research Objectives

This research focuses on consumers' comprehension of the verbal health warning in TV DTC advertisements (Kaphingst et al. 2005). In the first phase of the research, a detailed content analysis of a sample of the top advertised DTC prescription drugs provides firm answers about the executional format of current DTC commercials by documenting the

nature and valence of facial expressions during those commercials and how they vary by acts in the commercial. While such content analyses serve to document the landscape of DTC advertising, they do not address how specific executional cues and patterns in advertisements affect consumer processing, resulting beliefs about the advertised drug, or patients' understanding of important risks, all of which are relevant concerns to regulatory authorities and medical practitioners who ultimately prescribe these drugs.

The second objective in this research is to assess whether the valence of facial expressions during the warning segment of DTC commercials affects cognitive load, and, in turn, comprehension of the verbal health warning. This second phase of the research involves an eye-tracking (ET) experiment to measure cognitive load during the presentation of positive facial expressions in conjunction with a verbal health warning and to test the proposed relationships between altered cognitive processing of the warning and reduced consumers' understanding of a drug's risks as depicted in Figure 1.

More specifically, the study tests whether incongruence between the valence of characters' facial expressions during the health warning of DTC commercials affects the level of comprehension of the risks contained in the verbal health warning. The study uses a control baseline of neutral character facial expressions and proposes that, compared to neutral facial expressions, positive facial expressions presented in conjunction with a verbal health warning lead to greater cognitive processing load due to cross-modality incongruence and in turn lesser comprehension of the health warning. This proposed mediation model is summarized in the following hypotheses:

H1: Compared to neutral facial expressions, positive facial expressions during the health-warning act of DTC advertisements leads to reduced comprehension of the health warning.

It is posited that the lesser comprehension is due to the greater cognitive load created by the interference of positive visuals with negative risk information. The attentional capacity load is measured using pupil dilation, as assessed with ET, which extant pupillometry research has shown is sensitive to the level of cue congruency (Laeng et al. 2011; Laeng, Sirois, and Gredebäck 2012). In other words:

H2: Compared to neutral facial expressions, positive facial expressions during the health warning act of DTC advertisements increase cognitive load, as revealed by pupil dilation.

In turn, the increased cognitive load results in lower comprehension of the audio health warning in a mediated relationship, as hypothesized:

H3: The impact of the positive facial expressions on lower comprehension of the audio health warning is mediated through an increase in cognitive load.

The practical relevance of these hypotheses is grounded in the natural tension between advertisers' desire to hold viewers' attention and the government's intent to promote public understanding of product risk information. This research also investigates whether the addition of visual text, which is congruent with the audio risk information, can mitigate the distraction effect of the positive (incongruent) facial images. Currently, superimposed text

banners appear to be the FDA's primary tactic for mitigating presumed distraction effects in direct-to-consumer advertising due to positive visuals (Food and Drug Administration 2010). Beyond the regulatory relevance, the banner effect is studied as it may serve as a practical and conceptual marker for comparing the impact of various distractors in DTC advertising research. As documented in prior experiments testing subtitles in TV commercials, banners should redirect viewers' attention to the text and improve understanding of the verbal messages (Brasel and Gips 2014).

The final hypothesis proposes that including static text banners that are congruent with the message points of the health warning can mitigate the hypothesized decline in the level of comprehension of the health warning posited in H1-3. Text banners may sufficiently redirect attention away from positive facial expressions and, because the text is valence-congruent with the audio message, cognitive load should be reduced and message comprehension enhanced. Thus, it is hypothesized that:

H4: The presence of visual banners conveying the verbal information reduces the cognitive load created by valence-incongruent visuals and thereby improves comprehension of the health warning..

Study 1: Content Analysis

This first phase of this research is a detailed content analysis, a methodology widely used in analyzing media messages and advertising in particular. Content analysis provides a "scientific, objective, systematic, quantitative and generalizable description of communications content" (Kassarjian 1977, 10) and is useful in objectively documenting the media landscape (Berelson 1971; Kolbe and Burnett 1991). Its goal is descriptive: to identify patterns and frequencies of occurrences (Carlson 2008). In advertising research, it is used to document the nature and prevalence of executional formats, rhetorical figures, or message appeals (e.g., van Enschot and Hoeken 2015; Okazaki, Mueller, and Taylor 2010; Roberts and Pettigrew 2007). Many have conducted content analyses of DTC advertisements in magazines (Frosch et al. 2007) or websites (Macias and Lewis 2004), but to date no analysis has focused on facial expressions.

This study's content analysis follows established film theory guidelines: a shot-level cinematic analysis of the commercials' content with both auditory and visual elements of each shot coded (Sharff 1982). In film theory, a shot is a short (usually a few seconds) section filmed in a single running of the camera; in other words, when the camera angle changes, a new shot begins. Shot analysis captures the underlying structure of cinematic language; the shot, a crucial grammatical element of film, is used here to establish the syntax of the commercials (Sharff 1982). To assess differences in the portrayals of users by act, the nature and valence of facial expressions of the characters within each shot are compared across the three acts of the commercial: exposition (portion where the drug information is presented and its benefits), warning (portion where the risks are presented, usually in narrated form), and conclusion.

Sample

The analysis was conducted on a sample of 46 different TV commercials, including at least one advert from each of the top advertised DTC prescription drugs on US TV in 2011-12. These commercials were identified and downloaded using Shadow TV, a media-monitoring database allows a systematic and comprehensive search of the transcripts of all 210 US TV stations (network and cable). Analyses were conducted at the shot level.

Most DTC advertisements follow a similar format with three acts: 1) presentation of the benefits in Act 1; 2) presentation of the warning in Act 2; and 3) general information in Act 3. On average, the TV advertisements in the sample present 2.35 benefits, orally, compared to 13.02 side effects orally. In all advertisements, the verbal health warning (i.e., the major statement) was narrated (i.e., voice over; not by an on-screen actor). Off-screen narration is a central difference between the health warning segment (86.4% of the shots therein are narrated) and the remainder of the commercial (60.6% narrated; $X^2 (p < .01)$). No difference in the gender of the narrator was observed.

The dataset included 967 shots, with an average length of 3.26 seconds ($SD = 2.46$). Of the shots, 50.5% contained facial expressions. The exposition, warning and conclusion acts contained 441, 386 and 138 total shots respectively, and the analysis was based on the sample of the 486 shots with facial expressions: 234, 189, and 63 from each act respectively.

Coding procedures

Four coders were trained to use the well-established Facial Action Coding System (Ekman and Friesen 1978) to systematically categorize and codify microfacial expressions. Upon codification, the microfacial expressions can then be correlated to specific emotions (Boucher and Ekman 1975; Ekman and Rosenberg 2005). One of the authors attended an official FACS training course and then developed a study-specific version of the FACS manual, with examples from a practice set of DTC commercials, to serve as a reference for coding the commercials. To assist in categorizing a facial expression as positive or negative based on the microfacial features, the coding manual included a list of affect descriptors from previous research to aid in determining the valence (Russell 1980; Watson and Tellegen 1985). Examples of these descriptors are: delighted, glad, or cheerful (for positive) and distressed, annoyed, or alarmed (for negative).

To ensure accurate and reliable coding, the coders were trained for a period of four weeks using commercials not included in the sample. The phases of training included an orientation phase where coders reviewed the coding manual, procedures, and rules, and where they were introduced to the electronic coding forms; a group practice coding phase where they worked in dyads so they could compare their ratings and observe and discuss one another's coding abilities; and an individual practice phase where coders rated shots individually and then discussed the results in group meetings. Data collected during the individual practice phase were assessed for reliability using the Intraclass Correlation Coefficient (ICC). This statistic controls for agreement by chance that can inflate apparent consistency between coders (Krippendorff 1980). All of the measured ICCs were above the accepted level of .80 (Kassarjian 1977). If reliability was poor (below .70 on more complex

items or below .85 on less complex items), remedial group practice sessions were conducted until the acceptable level is achieved. When an acceptable level of consistency in coding was achieved, coding on the primary sample began.

Each shot with a facial expression was coded as positive, neutral, or negative. Two coders, blind to the research questions, coded the entire set of shots. To ensure no bias, shots containing facial expressions were coded, without audio information and in random order by two additional coders. Inter-coder reliability was 85.1%. Reconciliation was done by discussion.

Findings

Differences in facial expression valence between acts were assessed via Chi-square tests of proportions. The valence of characters' facial expressions differs significantly by act ($\chi^2(4) = 68.08, p < .01$): shots during the warning act display characters with positive facial expressions 69.8% of the time, neutral expressions 26.5% of the time and negative only 3.7% of the time. In contrast, during the exposition act, facial expressions are positive 47.0% of the time, neutral 26.1% of the time and negative 26.9% of the time. The conclusion act contained few facial expressions but 87.3% were positive, 9.5% negative and 3.2% neutral. The standardized residuals in the Chi-square test show that the proportion of positive facial expressions is significantly higher in the warning and conclusion act than in the exposition act and that, contrarily, the proportion of negative facial expressions is greater in the exposition act than in the other two acts. In other words, negative facial expressions are about 7 times less likely during the health warning and positive facial expressions are almost 1.5 as likely in the warning act than in the introductory act of the commercial. Appendix 1 contains a representative sample of the facial expressions encountered during the warning act.

Discussion

Study 1 shows that DTC advertisements present positively laden visual cues during the verbal conveyance of health warnings. As documented in previous research on responses to facial expressions in advertising (Lee, Kim, and Peng 2013; Xiao and Ding 2014) these positive facial cues communicate an overall positive message about the advertised product. The simultaneous presentation of a positive visual with the negatively valenced audio warning may create a more difficult processing environment due to cross-modality interference. As posited in the conceptual model this interference may ultimately prevents consumers from processing and comprehending the audio warning information.

Study 2: Eye-Tracking Experiment

To test whether positively valenced faces distract from and alter comprehension of the verbal message, Study 2 reports a controlled experiment that manipulates characters' facial expressions during the warning act to be either positive or neutral and that measures consumers' attention to and processing of the audio and visual information. An ET procedure was designed to assess whether the emotional facial expressions of on-screen characters during the health warning affect the level of health warning comprehension.

Methods

Study 2 was a controlled experiment with specially developed TV advertisements for a fictitious weight management drug. To test whether less congruent (i.e., positive) facial expressions in DTC advertisements create greater interference with processing of audio warning information than more congruent (i.e., neutral) facial expressions. Young female adult participants were randomly assigned to different commercial conditions according to a three-condition between-subjects experimental design that varied the valence of characters' facial expressions during the health warning of the commercial (to test H1-3) and the presence of a visual banner (to test H4). A two-step random allocation process was followed: first, one of the research assistants generated an unpredictable random sequence with Microsoft Excel, and this sequence was then used to assign each time slot with a scheduled participant to one of the three conditions; this process was conducted before the participants entered the lab, thus the nature of the treatment was concealed to all participants. The study was conducted in a university ET lab to provide pupillometry responses to the commercials (and test the proposed mediation model). The study was approved by the authors' Institutional Review Board.

Pharmaceutical drug and population context—A weight control medication was selected as relevant for the study population of young adult females. College-aged students, especially females, are particularly concerned with body image and weight management and many suffer from eating disorders (Eisenberg et al. 2011; Goswami, Sachdeva, and Sachdeva 2012). For example, one study found that approximately 74% of the normal-weight women report persistent or frequent concern with their body image (Brown University 2015).

Experimental stimuli—Three 90-second professional-quality commercials were filmed on location by a professional director and crew. Two female professional actors, one African-American and one Caucasian, were selected from a casting call based on their ability to convey authentic facial expressions and based on having a physique appropriate for a weight control commercial. Overall, the experimental commercials mimicked professional DTC advertisements in both quality and content: For instance, a female professional narrator did the voice-over for the health warning (30 seconds), and graphic artists produced all necessary graphic elements including the brand logo and text banners which synced with each audio warning message point in the banner condition (see below).

The commercial introduced *Oncazil* as a monthly prescription treatment that enables women with a BMI of over 25 to “lose at least 5 pounds per month with proper diet and exercise.” The health warning included a variety of common and severe adverse effects (e.g., temporary constipation; women who smoke may have higher risk of colon cancer). The message points and script for this fictitious brand were based upon previous FDA research (Aikin et al. 2011).

Manipulations: Facial expressions and banner—The audio, as well as the visuals in the exposition and conclusion acts (50 seconds and 10 seconds respectively), were constant across all three commercial conditions. The visuals for the 30-second health warning of the commercial were manipulated to create three conditions. In the two “No-Banner”

conditions, the commercials were identical except for the actors' facial expressions during the narrator's voice-over (which described the drug's health warning). In the "Positive No-Banner" condition both actors conversed – silently, of course - while exhibiting normal smiles. In the "Neutral No-Banner" condition, both actors exhibited normal, but no-smile, expressions during their silent dialogue. See Appendix 2 for sample visuals from each of the three conditions.

The conversation scenes switched between the actors, but only one actor appeared on screen. Thus, the "Positive-No Banner" and "Neutral-No Banner" conditions provide a test of H1-3. The "Positive-Banner" was identical to the "Positive No-Banner" except for a superimposed (static, not scrolling) text banner on the lower 20% of the screen. The banners summarized and synced with the narrator's audio warning points. Thus, the two "Positive" conditions provide a test of H4.

The face of each actor during the warning act was designated as the area of interest (AOI) for the ET analysis. The ASL D6 ET and software transforms the eye-gaze data into a series of variables for the AOI: the total amount of time (dwell duration; milliseconds), the number of eye fixations (count) and average pupil diameter (ET units; Applied Science Laboratories 2015). Pupil height and width were measured with better than .01 mm resolution with no averaging. The ET software fits an ellipse to the pupil in real time, for each sample (every 16.67 ms) outputting the length of the major axis (longest length, which at central fixation is the pupil width) and minor axis (shortest length, which at central fixation is the pupil height).

Statistical analysis—Analyses were performed using SPSS, version 22. The PROCESS Procedure for SPSS (Hayes 2013) was used to test whether pupil response mediates the hypothesized effects of facial expression and banner on health warning comprehension (Model 4 with bootstrap $N = 5,000$). A power analysis was performed for each two-cell contrast, based on a two-tailed test ($\alpha = .05$) with 80% probability to detect a difference in means of $d = .5$ (Faul et al. 2007; 2009). For a 12-item, dichotomous, objective comprehension measure, assuming probability of success on each item at .5 and a variance of 3 (i.e., $\sigma^2 = 12 \times .5 \times .5$), this difference translates into a mean difference of 0.87 and a necessary cell size of 64. However, some degree of attrition is assumed due to inability to calibrate on the ET (15% based on previous research; Brasel and Gips 2014). This results in a target cell size of at least 75 (i.e., $64 / .85$) or an N of 225. For a one-tailed test with $\alpha = .05$ this cell size provides an 80% probability to detect a difference in means of $d = .44$ (i.e., mean difference of 0.76.).

Sample—Female university students ($N = 223$) volunteered to participate in the study and were compensated for their time with credits in courses requiring participation in university-based research studies. The study was conducted in a university ET lab over a ten-month period. Participant ages ranged from 20 to 33 ($M = 21.5$, $SD = 1.4$). Caucasians were 62% of the sample and 7% were African American.

Their BMI, based on self-reported height and weight, ranged from 14 to 35 ($M = 22.5$, $SD = 3.2$). Table 1 provides means and standard deviations of all variables per condition and

shows whether contrasts between conditions are significant. Confirming the relevance of the selected drug for this population, 25.9% of the participants indicated they were currently on a prescription drug for weight management. Candidates for weight-loss drugs often have a BMI greater than 30 (Mayo Clinic 2015); yet there was no difference in our sample between those who reported currently being on a prescription for a weight management drug and those who were not: the BMIs were comparable ($M = 22.8$, $SD = 3.4$ vs. $M = 22.4$, $SD = 3.2$; $t(213) = .69$, $p > .05$). It is noteworthy that neither BMI nor prescription drug use was correlated with the comprehension dependent measure nor any of the ET variables and were therefore not included in the analyses.

Procedures—Individual sessions were held in the ET lab with controlled seating and lighting. The lab technician followed a standard orientation procedure: A brief explanation of the study process and review of the seating and equipment (ear-phone headset to prevent audio distractions, keyboard, and comfortable distance to the monitor with a mounted and unobtrusive ET camera).

The ASL optical/camera tracks the gaze of participants. A participant can sit freely in front of the monitor and move approximately one square foot. The lab technician was seated behind a screen as to provide anonymity for the participant during viewing and completion of the questionnaire. After the participant signed the informed consent form, a standard calibration process was conducted without making the participant aware of whether the calibration was actually successful. Each participant viewed an on-screen introduction which was also read aloud by a research assistant stating, "You will watch two advertisements about two new products, the first one is named Oncasil and the second is Airwalk. Please watch these advertisements as you normally would, after you watch them we will ask you a series of questions." Participants were randomly assigned to one of the three 90-second Oncasil commercial conditions, which was then followed by the Airwalk advertisement. This second advertisement, unrelated to the main study, served as a short distractor between viewing the DTC advertisement and completing the outcome measures.

Participants then completed a self-paced online questionnaire. Participants first answered several questions about the filler advertisement and then standard 5-point semantic differential scale items measuring attitude toward the Oncasil advert (Att_{ad} ; 4 items; e.g., Very Bad – Very Good; Cronbach's $\alpha = .87$) and brand (Att_{br} ; 3 items; $\alpha = .87$). This was followed by the warning comprehension questions and several manipulation check items. Participants were asked to indicate, "how the women in the commercial seemed" on several characteristics, including two semantic differential scales for valence (very unhappy (1) - very happy (5); negative (1) – positive (5); $\alpha = .72$). Sessions lasted an average of 12 minutes.

Dependent measure—Few studies have assessed consumer comprehension of DTC TV advertising, and those that exist are limited in scope. For instance, one study examined individual characteristics related to viewer recall of information contained in a single DTC commercial and another examined the relationship between literacy and comprehension of various message points contained in three DTC commercials (Kaphingst et al. 2005; Schommer, Doucette and Mehta 1998). In this study, the warning comprehension measure

was the number of correct responses to twelve objective items (six true and six false), based on the audio message points during the warning act (see items in Appendix 3). This measure is consistent with extant research on consumers' objective comprehension of marketing communications in general and on existing measures of understanding of risk in particular (Mick 1992; Pillai, Katsikeas and Presi 2012; Weinstein 1999). These true/false tests tap specific elements of the communicated messages to assess their ability to be understood and are common ways to assess the value of health information (Whittingham et al. 2008).

Results

The cell sizes for the Positive-No Banner (PNB), Neutral-No Banner (NNB) and Positive-Banner (PB) were 75, 73, and 75 respectively. Seven participants did not complete the warning comprehension questions and these were omitted from the analyses (net cell sizes of 73, 69, and 74 ($n = 216$)). Means per condition are reported in Table 1, and correlations between outcome variables are in Table 2.

Manipulation checks—As expected, participants rated the women as more positive in the positive no-banner condition than in the neutral condition (4.01 vs. 3.76; $t(213) = 2.07, p < .05$). In contrast, ratings on other characteristics did not differ (e.g., seriousness ($t(213) = 1.54; p = .12$), maturity ($t(213) = 0.49; p = .62$)). Furthermore, the manipulations did not affect participants' overall attitude toward the ad ($F(2, 213) < 1; p = .75$) or the brand ($F(2, 213) < 1; p = .56$). Together, these checks suggest that the actors' facial expressions successfully changed affective perceptions of the actors without creating an unnatural impression or confounded treatment.

Health warning comprehension—Comprehension varied as hypothesized across the three conditions (ANOVA $F(2, 213) = 4.12, p < .05; \eta^2 = .037$; 95% CI: .001 – .09). In line with H1, the first planned contrast confirmed that comprehension scores were lower amongst participants in the positive facial condition than those in the neutral condition (6.98 vs. 7.42; $t(121) = -1.70, p = .046$ (one-tail)). The second planned contrast reveals that, as hypothesized by H4, the addition of a banner to the positive facial commercial increased comprehension (7.78 vs. 6.98; $t(122) = 2.93, p < .01$). Here, and where appropriate, contrast tests are reported not assuming equal variances when indicated by a significant Levene's test ($2, 213) = 9.30, p < .001$).

Face fixation counts and duration of face dwell—Calibration of eye gaze was achieved for 189 (87.5%) of participants. Thus, the cell sizes for analyses involving ET measures are: 62 (PNB), 62 (NNB), and 65 (PB). The pattern of facial fixations during the warning act varied significantly across conditions ($F(2, 186) = 11.57, p < .001; \eta^2 = .11$; 95% CI: .03 – .17). The contrasts reveal how the banner changes the pattern of focal attention. Participants in the positive and neutral facial conditions exhibited similar counts (10.2 vs. 10.7; $t(122) = -.46; p = .65$). As expected, the addition of a banner in the positive facial condition dramatically reduced number of fixations on the AOI (10.2 vs. 6.08; $t(122) = -3.88, p < .001$).

The pattern for face dwells matched that of the fixation counts and varied significantly across the experimental conditions ($F(2, 186) = 19.60, p < .001; \eta^2 = .17; 95\% \text{ CI: } .07 - .24$). Participants in the positive and neutral conditions spent the same amount of time on the faces (10.4 vs. 9.1 seconds; $t(122) = 1.19; p = .23$) but the addition of a banner in the positive facial condition dramatically reduced dwell time (10.4 vs. 4.5; $t(123) = 5.95, p < .001$).

Average pupil diameter—Pupil responses were examined for those participants who fixated on the AOI during the warning act (i.e., facial fixation count > 0 ; $n = 181$). The cell sizes for analyses involving pupil responses are: 60 (PNB), 60 (NNB), and 61 (PB).

The pattern of means for average pupil diameter (APD) was the inverse of comprehension. While the overall F test for the three conditions was not statistically significant ($F(2, 178) = 2.69, p = .07$), the specific contrast pattern of interest was ($t(178) = 2.28; p < .05$): participants in the positive facial conditions exhibited greater dilation than those in the neutral condition (49.32 vs. 46.11; $t(115) = 1.95, p < .05$ (one-tail); 90% CI: $-6.20 - -.22$; $d = .36$; 95% CI: $-.01 - .71$). As expected the addition of a banner in the positive facial condition dramatically reduced dilation (49.32 vs. 45.4; $t(109) = 2.17, p < .05$ (one-tail); 90% CI: $-6.92, -.96$; $d = .39$; 95% CI: $.03 - .75$).

Mediation analysis—H2 predicted that processing and encoding of audio warning (negative) information is disrupted more by cues that are less congruent in valence (positive facial expressions) than by those that are more congruent (neutral facial cues), and this process may be indicated by pupil responses. To further test the theoretical model and H3, the mediating role of APD was tested in a mediation analysis (Preacher, Rucker, and Hayes 2007). The two no-banner conditions were used to estimate the indirect effect of valence (see Table 3 for path coefficients). This analysis reveals a statistically significant mediating effect of APD for the relationship between facial expression valence and health warning comprehension ($\beta_{\text{Valence} - \text{APD} - \text{Comprehension}} = .109 (.093)$). Furthermore, as reported in Table 3, when APD is included as a mediator, the direct effect of valence is no longer statistically significant ($p = .238$). This supports the hypothesized mediation process and suggests that an alternative or “omitted” mediator is unlikely (Zhao, Lynch, and Chen 2010).

The two positive conditions were used to estimate the mediator role of APD for the effect of banner on comprehension. Again, the analysis reveals a statistically significant mediating effect of APD for the relationship between the presence of a banner and health warning comprehension ($\beta_{\text{Banner-APD-Comprehension}} = 0.170 (.120)$). When APD is included in the analysis, as reported in Table 3, the direct effect of banner (.437) is not statistically significant ($p = .15$), again providing support for the mediating role of APD and H3. Similar analyses show that fixation count and dwell do not mediate the effects of valence or banner.

General Discussion

Grounded in the context of DTC commercials, this research empirically demonstrates that characters with positive expressions interfere with comprehension of the concurrent verbal warning message. The ET data and results presented here are consistent with emerging

research on the importance of cue congruency during dual modality processing (Kuchinsky et al. 2013, 2014). Processing and encoding of the audio warning information is disrupted more by cues that are less congruent in valence (positive facial expressions) than those that are more congruent (neutral facial cues).

Theoretical implications

This study adds to the current body of research on the processing and impact of facial imagery, this research shows that even subtle facial expressions of characters in advertisements can affect the processing of simultaneously presented auditory information (Small and Verrochi 2009; Tanner and Maeng 2012). The findings show that incongruence in the valence of the visual and the auditory information leads to an increase in cognitive load and poorer understanding of the verbal message. Given that previous research has found that positively valenced faces often accompany serious messages, such as breast cancer information (McWirth et al. 2012), this finding suggests that closer attention should be paid to the subtle elements of visual rhetoric in advertising and to the fit between those visual cues and the content of the accompanying texts or verbal messages. This research contributes to the growing body of evidence that the interference between images and text can severely bias consumers' responses to and understanding of the information (e.g., Parguel, Benoît-Moreau, and Russell 2015)

Another contribution of this research is in the use of ET technology. This non-invasive but reliable tool for assessing attention to visual elements is promising for measuring attention to and processing of both print and audiovisual advertising messages. The ability to identify areas of interests within a static or moving image offers tremendous potential for assessing the relative attention to different visual elements of an advertisement and for corroborating these data with advertising outcomes, such as memory, attitude, or comprehension (Pieters and Wedel 2004; Brasel and Gips 2014). Pupillometry offers novel insights into cognitive load, another black-box issue of information processing. In particular, the findings show that pupil dilation is a useful measure of cognitive effort in response to stimuli in general (Allard, Wadlinger, and Isaacowitz 2010) and a measure of listening difficulty in particular (Kuchinsky et al. 2014).

Policy implications

Concerns surrounding the distractors during the mandatory health warning of DTC advertisements persist and are fueled by the consistent findings of content analyses. In the US, the debate has received renewed attention following the American Medical Association's (2015) recent call for the federal government to outright prohibit DTC pharmaceutical advertising. The Association expresses concerns that the practice drives consumer demand for inefficacious treatments and contributes to ever-increasing healthcare costs. In fact, prescription drug costs are among the top healthcare problems in the US (Kaiser Family Foundation 2015).

This study, designed with a commercial containing fewer and less dramatic visuals and scene changes than those normally found in DTC TV commercials, offers a conservative test of the degree to which positive visuals detract from important verbal information about drug

risks. Yet, the findings clearly show that congruency of a specific visual element impacts audio processing, with incongruence, in the form of positive visuals presented during the warning segment, impairing comprehension of the health warning. This manipulation was done while not significantly affecting viewers' overall affective feelings toward either the brand or the advertisement, two criteria of concern to advertisers.

The findings also clarify the potential mitigation role of well-designed banners and suggest that simply requiring a banner during the warning, as currently proposed by the FDA, will not likely enhance comprehension *unless* such banners refocus focal attention *and* do so in a manner that reduces overall cognitive load during the warning (Food and Drug Administration 2010). Banners designed to be congruent with audio warning information can enhance comprehension and possibly reduce distraction effects such as the positive facial expressions used in this experiment. This study's findings suggest that focal attention must be redirected from the distractor elements and is a necessary requirement before an ad can effectively convey a drug's risk. Thus, the findings may provide some direction for how policymakers and sponsors can use objective data to define "clear, conspicuous and neutral" executions of the risk statement, as desired by the FDA.

More generally, these studies extend previous non-experimental research on dual-modality effects on viewer comprehension, which addressed the ability of concurrent audio and video to enhance comprehension (Murray, Manrai, and Manrai 1998). This research signals key weaknesses in policies that merely require text or verbal warnings and disclosures in advertising to correct consumers' judgments or understanding of a message. Careful attention must be paid to the potential for cross-modality interference and the burden it creates on cognitive processing and understanding of the message. This concern is especially salient with vulnerable audiences, such as children or the elderly, who may be especially susceptible to the influence of visuals and whose cognitive resources may be overly tasked in a multimodality environment. Finally, this research highlights the ethical concern of advertising drugs directly to consumers instead of leaving those decisions to the medical professionals.

Limitations and future research

While the study offers solid scientific evidence of visual distraction, some limitations exist and present opportunities for future research. The single exposure context in a controlled environment offered great internal validity for the ET study but also calls for replications and extensions in other contexts and with a broader range of stimuli. Moderators of the effect identified here could be tested; for instance, future research could incorporate the moderating roles of literacy or other individual variables likely to affect attention to the visuals and comprehension (Kaphingst et al. 2005; Schommer, Doucette and Mehta 1998).

Future research could also assess whether, as in other advertising contexts, the impact of positive visuals lessens with increased exposure and what, in turn, may be the consequences on cross-modality interference and message comprehension. While this study's focus on young adult women was appropriate for the prescription drug tested, future studies should be conducted with a broader range of demographic segments of the population targeted by DTC advertisements or with products other than pharmaceutical drugs. Lastly, turning the tables

as it were, coupling the ET methodology with new technologies that can read and emotionally categorize a TV viewer's facial expression when watching an advertisement might provide more robust insights into how cross-modality interferences manifest (Lewinski, Fransen, and Tan 2014). Although the FACS is a widely used and is informative to document the nature of advertisements' content, as conducted in Study 1, it may lack enough sensitivity to measure emotional reactions to advertisements (Derbaix 1995). Relying on automated physiological measures would provide more objective assessment of consumers' responses to the facial expressions and could capture their emotional responses (Poels and Dewitte 2006).

In conclusion, this study offers empirical evidence of cross-modal interference: incongruent visuals can impair understanding of auditory information in commercials. Given that the visual distraction effect emerged with a relatively subtle manipulation of neutral versus positive facial expressions, the findings further signal that the issue of visual distraction would hold and perhaps worsen with more complex visuals (e.g., more engaging or multiple visual cues). Increasing the cognitive load for viewers is unlikely to improve comprehension of a verbal message. Instead, it is more likely to further hinder comprehension.

References

- Aikin, Kathryn J., O'Donoghue, Amie C., Swasy, John L., Sullivan, Helen W. A randomized trial of risk information formats in direct-to-consumer prescription drug advertisements. *Medical Decision Making*. 2011; 31(6):23–33.
- Allard, Eric S., Wadlinger, Heather A., Isaacowitz, DM. Positive gaze preferences in older adults: Assessing the role of cognitive effort with pupil dilation. *Aging Neuropsychology and Cognition*. 2010; 17:296–311.
- American Medical Association. AMA calls for ban on direct to consumer advertising of prescription drugs and medical devices. 2015. <http://www.ama-assn.org/ama/pub/news/news/2015/2015-11-17-ban-consumer-prescription-drug-advertising.page>
- Amyx, Douglas, Bristow, Dennis, Robb, Jeffrey. Unintended effects of affirmative disclosures in print advertisements. *Journal of Legal, Ethical and Regulatory Issues*. 2004; 7(1) 89-10.
- Applied Science Laboratories. D6 desktop eye tracking. 2015. <http://www.asleynettracking.com/Site/Portals/0/D6.pdf>
- Batra, Rajeev, Ray, Michael L. Affective responses mediating acceptance of advertising. *Journal of Consumer Research*. 1986; 13:234–249.
- Berelson, B. Content analysis in communication research. Glencoe: Free Press; 1971.
- Boucher, Jerry D., Ekman, Paul. Facial areas and emotional information. *Journal of Communication*. 1975; 25(2):21–29. [PubMed: 1127138]
- Brasel, Adam S., Gips, James. Enhancing television advertising: Same-language subtitles can improve brand recall, verbal memory, and behavioral intent. *Journal of the Academy of Marketing Science*. 2014; 42(3):322–336.
- Brown University Health Education Department. Body Image. 2015. http://www.brown.edu/Student_Services/Health_Services/Health_Education/nutrition_&_eating_concerns/body_image.php
- Carlson, Les. Use, misuse, and abuse of content analysis for research on the consumer interest. *The Journal of Consumer Affairs*. 2008; 42(1):100–105.
- Castendyk, Oliver, Dommering, Egbert J. European media law. The Netherlands: Kluwer Law International; 2008.
- De Gelder, Beatrice, Vroomen, Jean. The perception of emotion by ear and by eye. *Cognitive Emotion*. 2000; 14:289–311.

- Derbaix, Christian M. The impact of affective reactions on attitudes toward the advertisement and the brand: a step toward ecological validity. *Journal of Marketing Research*. 1995; 32(4):470–479.
- Donohue, Julie M., Cevalco, Marisa, Rosenthal, Meredith B. A decade of direct-to consumer advertising of prescription drugs. *The New England Journal of Medicine*. 2007; 357(7):673–681. [PubMed: 17699817]
- Donohue SE, Appelbaum LG, Park CJ, Roberts KC, Woldorff MG. Cross-modal stimulus conflict: The behavioral effects of stimulus input timing in a visual-auditory Stroop task. *PLoS ONE*. 2013; 8(4):e62802. [PubMed: 23638149]
- Eisenberg, Daniel, Nicklett, Emily J., Roeder, Kathryn, Kirz, Nina E. Eating disorder symptoms among college students: Prevalence, persistence, correlates, and treatment seeking. *Journal of American College Health*. 2011; 59(8):700–707. [PubMed: 21950250]
- Ekman, Paul, Friesen, Wallace V. *The facial action code: A manual for the measurement of facial movement*. Palo Alto: Consulting Psychologist Press; 1978.
- Ekman, Paul, Rosenberg, E., editors. *What the face reveals*. 2. New York: Oxford University Press; 2005.
- Engelhardt, Paul E., Ferreira, Fernanda, Patsenko, Elena G. Pupillometry reveals processing load during spoken language comprehension. *The Quarterly Journal of Experimental Psychology*. 2010; 63(4):639–645. [PubMed: 20017058]
- van Enschoot, Renske, Hoeken, Hans. The occurrence and effects of verbal and visual anchoring of tropes on the perceived comprehensibility and liking of TV commercials. *Journal of Advertising*. 2015; 44(1):25–36.
- Etcoff, Nancy L., Magee, John J. Categorical perception of facial expressions. *Cognition*. 1992; 44:227–240. [PubMed: 1424493]
- Faul, Franz, Erdfelder, Edgar, Lang, Albert-Georg, Buchner, Alex. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*. 2007; 39:175–191. [PubMed: 17695343]
- Faul, Franz, Erdfelder, Edgar, Buchner, Alex, Lang, Albert-Georg. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*. 2009; 41:1149–1160. [PubMed: 19897823]
- Food and Drug Administration. Direct-to-consumer prescription drug advertisements: Presentation of the major statement in television and radio advertisements in a clear, conspicuous, and neutral manner: proposed rule. *Federal Register*. 2010; 75(59):5376–5387.
- Frosch, Dominick L., Krueger, Patrick M., Hornik, Robert C., Cronholm, Peter F., Barg, Francis K. Creating demand for prescription drugs: A content analysis of television direct-to-consumer advertising. *Annals of Family Medicine*. 2007; 5(1):6–13. [PubMed: 17261859]
- Goswami, Shweta, Sachdeva, Sandeep, Sachdeva, Ruchi. Body image satisfaction among female college students. *Industrial Psychiatry Journal*. 2012; 21(2):168–172. [PubMed: 24250055]
- Hareli, Shlomo, Shomrat, Noga, Hess, Ursula. Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion*. 2009; 9(3):378–384. [PubMed: 19485615]
- Hayes, Andrew F. *Introduction to mediation, moderation and conditional process analysis*. New York, NY: Guilford Press; 2013.
- Hertrich I, Dietrich S, Ackermann H. Cross-modal interactions during perception of audiovisual speech and nonspeech signals: An fMRI study. *Journal of Cognitive Neuroscience*. 2011; 23:221–237. [PubMed: 20044895]
- Hollon, Matthew F. Direct-to-consumer advertising: A haphazard approach to health promotion. *JAMA*. 2005; 293(16):2030–2033. [PubMed: 15855439]
- Hoy MG, Andrews JC. Adherence of prime-time televised advertising disclosures to the "clear and conspicuous" standard: 1990 versus 2002. *Journal of Public Policy & Marketing*. 2004; 23(2):170–182.
- Hyönä, Jukka, Tommola, Jorma, Alaja, Anna-Mari. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology*. 1995; 48A:598–612.
- Johar, Gita V., Simmons, Carolyn J. The use of concurrent disclosures to correct invalid inferences. *Journal of Consumer Research*. 2000; 26(4):307–322.

- Kahneman, Daniel. Attention and effort. Englewood Cliffs, NJ: Prentice-Hall; 1973.
- Kaiser Family Foundation. Prescription drug costs remain atop the public's national health care agenda, well ahead of Affordable Care Act revisions and repeal. 2015. <http://kff.org/health-costs/press-release/prescription-drug-costs-remain-atop-the-publics-national-health-care-agenda-well-ahead-of-affordable-care-act-revisions-and-repeal>
- Kantar Media. MARS OTC/DTC Consumer health study. New York, NY: 2014.
- Kaphingst, Kimberly A., Rudd, Rima E., DeJong, William, Daltroy, Lawren H. Comprehension of information in three direct-to-consumer television prescription drug advertisements among adults with limited literacy. *Journal of Health Communication*. 2005; 10(7):609–619. [PubMed: 16278198]
- Kassarjian, Harold H. Content analysis in consumer research. *Journal of Consumer Research*. 1977; 4:8–18.
- Koelewijn, Thomas, Zekveld, Adriana A., Festen, Joost M., Rönnerberg, Jerker, Kramer, Sophia E. Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology*. 2012
- Kolbe, Richard H., Burnett, Melissa S. Content-analysis research: An examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research*. 1991; 18:243–250.
- Kolbe, Richard H., Albanese, Paul J. Man to man: A content analysis of sole-male images in male-audience magazines. *Journal of Advertising*. 1996; 25(4):1–20.
- Krippendorff, Klaus. Content analysis: An introduction to its methodology. Beverly Hills: Sage Publications; 1980.
- Kuchinsky, Stefanie E., Ahlstrom, Jayne B., Vaden, Kenneth I., Cute, Stephanie L., Humes, Larry E., Dubno, Judy R., Eckert, Mark A. Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology*. 2013; 50(1):23–34. [PubMed: 23157603]
- Kuchinsky, Stefanie E., Ahlstrom, Jayne B., Cute, Stephanie L., Humes, Larry E., Dubno, Judy R., Eckert, Mark A. Speech-perception training for older adults with hearing loss impacts word recognition and effort. *Psychophysiology*. 2014; 51(10):1046–1057. [PubMed: 24909603]
- Laeng, Bruno, Ørbo, Marte, Holmlund, Terje, Miozzo, Michele. Pupillary stroop effects. *Cognitive Processing*. 2011; 12:13–21. [PubMed: 20865297]
- Laeng, Bruno, Sirois, Sylvain, Gredebäck, Gustaf. Pupillometry: A window to the preconscious. *Perspectives on Psychological Science*. 2012; 7(1):18–27. [PubMed: 26168419]
- Lee, Mira, Kim, Mikyoung, Peng, Wei. Consumer reviews: Reviewer avatar facial expression and review valence. *Internet Research*. 2013; 23(2):116–132.
- Lewinski, Peter, Fransen, Marieke L., Tan, EdSH. Predicting advertising effectiveness by facial expressions in response to amusing persuasive stimuli. *Journal of Neuroscience, Psychology, and Economics*. 2014; 7(1):1–14.
- Macias, Wendy, Lewis, Liza Stavchansky. A content analysis of direct-to-consumer (DTC) prescription drug web sites. *Journal of Advertising*. 2004; 32(4):43–56.
- Mack, John. Pharma promotional spending in 2013. *Pharma Marketing News*. 2014:13.
- Massaro, Dominic W., Egan, Peter B. Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*. 1996; 3:215–221. [PubMed: 24213870]
- Marks LE. On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*. 1987; 13:384–394. [PubMed: 2958587]
- Matusz PJ, Eimer M. Multisensory enhancement of attentional capture in visual search. *Psychonomic Bulletin & Review*. 2011; 18:904–909. [PubMed: 21748418]
- Mayo Clinic. Healthy Lifestyle - Weight Loss. 2015. <http://www.mayoclinic.org/healthy-lifestyle/weight-loss/in-depth/weight-loss-drugs/art-20044832>
- McQuarrie, Edward F., Phillips, Barbara J. Go figure! New directions in advertising rhetoric. New York, NY: Taylor and Francis; 2008.

- Mcwhirther JE, Hoffman-goetz L, Clarke JN. Can you see what they are saying? Breast cancer images and text in Canadian women's and fashion magazines. *Journal of Cancer Education*. 2012; 27(2): 383–391. [PubMed: 22228485]
- Mick, David Glen. Levels of subjective comprehension in advertising processing and their relations to ad perceptions, attitudes, and memory. *Journal of Consumer Research*. 1992; 18(4):411–424.
- Morgan, Fred W., Stoltman, Jeffrey J. Television advertising disclosures: An empirical assessment. *Journal of Business and Psychology*. 2002; 16(4):515–535.
- Murray, Noel M., Manrai, Lalita A., Manrai, Ajay K. How super are video supers? A test of communication efficacy. *Journal of Public Policy & Marketing*. 1998; 17(1):24–34.
- New Zealand Ministry of Health. Direct-to-consumer advertising of prescription medicines in New Zealand: Summary of submissions. Wellington: Ministry of Health; 2011.
- Okazaki, Shintaro, Mueller, Barbara, Taylor, Charles Ray. Measuring soft-sell versus hard-sell advertising appeals. *Journal of Advertising*. 2010; 39(2):5–20.
- Paivio, Alan. Imagery and verbal processes. New York: Holt, Rinehart, and Winston, Inc.; 1971.
- Parguel, Béatrice, Benoît-Moreau, Florence, Russell, Cristel A. Can nature-evoking elements in advertising greenwash consumers? The power of 'executional greenwashing.'. *International Journal of Advertising*. 2015; 34(1):107–134.
- Pieters, Rik, Wedel, Michel. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*. 2004; 68:36–50.
- Pillai, Kishore Gopalakrishna, Katsikeas, Constantine, Presi, Caterina. Print advertising: Type size effects. *Journal of Business Research*. 2012; 65(6):865–868.
- Poels, Karolien, Dewitte, Siegfried. How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research*. 2006; 46(1):18–37.
- Preacher, Kristopher J., Rucker, Derek D., Hayes, Andrew. Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*. 2007; 42:185–227. [PubMed: 26821081]
- Roberts, Michele, Pettigrew, Simone. A thematic content analysis of children's food advertising. *International Journal of Advertising: The Review of Marketing Communications*. 2007; 26:357–367.
- Russell, Cristel A. Investigating the effectiveness of product placements in television shows: The role of modality and plot connection congruence on brand memory and attitude. *Journal of Consumer Research*. 2002; 29(3):306–318.
- Russell, James. A circumplex model of affect. *Journal of Personality and Social Psychology*. 1980; 39(6):1161–1178.
- Schommer, Jon C., Doucette, William R., Mehta, Bella H. Rote learning after exposure to a direct-to-consumer television advertisement for a prescription drug. *Clinic Therapeutics*. 1998; 20(3):617–632.
- Scott, Linda M. Images in advertising: The need for a theory of visual rhetoric. *Journal of Consumer Research*. 1994; 21:252–273.
- Sengupta, Jaideep, Fitzsimmons, Gavin J. The effect of analyzing reasons on the stability of brand attitudes: A reconciliation of opposing predictions. *Journal of Consumer Research*. 2004; 31(3): 705–711.
- Sharff, Stefan. The elements of cinema: Toward a theory of cinesthetic impact. New York: Columbia University Press; 1982.
- Sirois, Sylvain, Brisson, Julie. Pupillometry. *Wiley Interdisciplinary Review: Cognitive Science*. 2014; 5(6):679–692.
- Small, Deborah A., Verrochi, Nicole M. The face of need: Facial emotion expression on charity advertisements. *Journal of Marketing Research*. 2009; 46:777–787.
- Spence, Charles. Crossmodal correspondences: A tutorial review. *Attention, Perception and Psychophysics*. 2011; 73(4):971–995. Retrieved from <http://search.proquest.com/docview/920261733?accountid=8285>.
- Tanner, Robin J., Maeng, Ahreum. A tiger and a president: Imperceptible celebrity facial cues influence trust and preference. *Journal of Consumer Research*. 2012; 39:769–83.

- Tavassoli, Nader T., Lee, Yih Hwai. The differential interaction of auditory and visual advertising elements with Chinese and English. *Journal of Marketing Research*. 2003; 40:468–480.
- van der Molen J. Assessing text-picture correspondence in television news: The development of a new coding scheme. *Journal Of Broadcasting & Electronic Media*. 2001; 45(3):483–498.
- Vatakis A, Spence C. Evaluating the influence of the "unity assumption" on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica*. 2008; 127:12–23. [PubMed: 17258164]
- Waldmann, Michael R., Holyoak, Keith J. Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*. 1992; 121:222–236. [PubMed: 1534834]
- Watson, David, Tellegen, Auke. Toward a consensual structure of mood. *Psychological Bulletin*. 1985; 98:219–235. [PubMed: 3901060]
- Weinstein, Noel. What does it mean to understand a risk? Evaluating risk comprehension. *Journal of the National Cancer Institute Monographs*. 1999; 25:15–20.
- Whittingham, Jill, Robert, RD., Ruiter, AC., Castermans, Diana, Huiberts, Annemarie, Kok, Gerjo. Designing effective health education materials: Experimental pre-testing of a theory-based brochure to increase knowledge. *Health Education Research*. 2008; 23:414–426. [PubMed: 17586587]
- Wolfe, Sidney M. Direct-to-consumer advertising - education or emotion promotion. *New England Journal of Medicine*. 2002; 346:524–526. [PubMed: 11844857]
- Xiao, Li, Ding, Min. Just the faces: Exploring the effects of facial features in print advertising. *Marketing Science*. 2014; 33(3):338–352. 459–460.
- Zekveld, Adriana A., Heslenfeld, Dirk J., Johnsrude, Ingrid S., Versfeld, Niek J., Kramer, Sophia E. The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*. 2014; 101:76–86. [PubMed: 24999040]
- Zekveld, Adriana A., Kramer, Sophia E. Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*. 2014; 51(3):277–284. [PubMed: 24506437]
- Zhao, Xinshu, Lynch, John G., Chen, Qimei. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*. 2010; 37(2):197–206.

Appendix 1: Sample Facial Expressions during the Health Warning Act of DTC Commercials in the Content Analysis

Appendix 2. Visual Examples from the DTC Commercial Stimuli in Study 2

Neutral facial expressions – Neutral No Banner (NNB)

Happy facial expressions – Positive No Banner (PNB)

Happy facial expressions – Positive Banner (PB)

Appendix 3. Health Warning Comprehension Items

“Based on your understanding of this commercial, which of the following are true?” (12 items presented in random order)

False items

Women who smoke while they take Oncasil increase their risk of high blood pressure.

Oncazol has not been shown to be habit-forming.

Oncazol may cause dry mouth.

Oncazol may increase the risk of high blood pressure.

Heart valve damage has been seen in rare cases.

Oncazol may cause temporary irritability.

True items

Your risk of colon cancer may increase if you take Oncazol and you smoke.

Women who are trying to become pregnant should not take Oncazol.

Oncazol should not be used with weight loss diet plans.

Oncazol may increase the risk of headaches.

Oncazol may increase the risk of urinary tract infections.

Possible side effects include constipation.

Figure 1.
Conceptual Model

Table 1

Means (Standard Deviations) By Condition In Study 2

	Variable	Neutral – No Banner	Positive – No Banner	Positive - Banner
	Cell size	69	73	74
Manipulation checks and controls	Use of prescription weight drug (%)	22.1	29.2	23.3
	BMI	22.27 ^{a b} (3.23)	23.22 ^a (3.61)	22.07 ^b (2.97)
	Valence perception of characters	3.76 ^b (.74)	4.01 ^a (.68)	3.71 ^b (.74)
	Attitude toward the Ad	2.31 ^a (.88)	2.46 ^a (.81)	2.36 ^a (.84)
	Attitude toward the Brand	2.72 ^a (.81)	2.80 ^a (.91)	2.69 ^a (.95)
Dependent Variable	Health Warning Comprehension	7.42 ^{a b} (1.75)	6.98 ^a (1.23)	7.78 ^b (1.98)
Eye-tracking measures (during warning)	Face Fixation Counts	10.7 ^a (6.46) (<i>n</i> = 62)	10.2 ^a (6.05) (<i>n</i> = 62)	6.08 ^b (5.44) (<i>n</i> = 65)
	Duration of Face Dwell	9.11 ^a (6.4)	10.44 ^a (6.01)	4.52 ^b (4.24)
	Average Pupil Diameter (APD)	46.11 ^{a b} (9.7) (<i>n</i> = 60)	49.32 ^a (8.2) (<i>n</i> = 60)	45.40 ^b (11.4) (<i>n</i> = 61)

Within each row, means with no common letter in the superscript are significantly different from each other for two-tailed tests at $p < .05$.

Table 2

Correlations between Outcome Variables (Study 2)

Outcome Variables ⁺				
Variable	Comprehension	Fixation Count	Dwell Time	A P D
Comprehension	7.29 (1.69)	-.16 *	-.17 *	-.24 **
Fixation Count	-1.64	9.36 (6.16)	.88 **	.22 **
Dwell Time	-1.73	32.90	8.32 (6.03)	.26 **
Average Pupil Diameter (APD)	-4.09	13.63	15.90	46.93 (10.00)

⁺Mean and standard deviation on diagonal with Pearson correlation above diagonal and covariance below ($n = 181$).

*Correlation significant at the .05 level (2-tailed).

**Correlation significant at the .01 level (2-tailed).

Pupil Dilation as Mediator of Facial Expression Valence and Banner Effect on Comprehension in Study 2

Table 3

Facial Expression Valence Path Model	Coef	SE	t	p	LLCI	ULCI
Valence* - APD	-3.21	1.65	-1.94	.054	-6.48	.053
APD - Comprehension	-.034	.015	-2.26	.025	-.063	-.004
Valence - Comprehension	.324	.273	1.18	.238	-.217	.865
Indirect Effect of Facial Expression Valence on Comprehension						
APD	Effect	Boot SE	Boot LLCI	Boot UCI		
	.109	.093	.004	.388		

Banner Path Model	Coef	SE	t	p	LLCI	ULCI
Banner** - APD	-3.93	1.81	-2.16	.032	-7.53	-0.34
APD - Comprehension	-.043	.015	-2.87	.005	-.073	-.013
Banner - Comprehension	.437	.304	1.43	.153	-.165	1.039
Indirect Effect of Banner on Comprehension						
APD	Effect	Boot SE	Boot LLCI	Boot UCI		
	.170	.120	.011	.508		

* Valence (Neutral, No-Banner = 1, Positive, No-Banner = 0)

** Banner (Positive, Banner = 1, Positive, No-Banner = 0)

A P D = Average Pupil Diameter

Process Procedure with Model 4, SPSS 2.13; bootstrap N=5,000.