

Special issue  
Credibility

**Jaime López Díez**

<https://orcid.org/0000-0003-3443-5910>  
jailop05@ucm.es  
Universidad Complutense de  
Madrid

**Jesús Bermejo-Berros**

<http://orcid.org/0000-0002-2830-604X>  
lipsimedia@gmail.com  
Universidad de Valladolid

**Submitted**

April 23th, 2018

**Approved**

September 26th, 2018

© 2019

Communication & Society  
ISSN 0214-0039  
E ISSN 2386-7876  
doi: 10.15581/003.32.1.91-106  
[www.communication-society.com](http://www.communication-society.com)

2019 – Vol. 32(1)

pp. 91-106

**How to cite this article:**

López Díez J. & Bermejo-Berros, J. (2019). Diegetic and non-diegetic surprises, and their effect on liking, long-term recall and comprehension in narrative television commercials. *Communication & Society*, 32(1), 91-106.

## Diegetic and non-diegetic surprises, and their effect on liking, long-term recall and comprehension in narrative television commercials

**Abstract**

The aim of this experiment is to classify surprises in audiovisual narratives, and to measure the efficacy of surprise in audiovisual stories in terms of liking, long-term recall and comprehension in television commercials. The theoretical analysis leads to distinguish 3 types of audiovisual narrative surprises: non-diegetic, diegetic *implausible* and diegetic *plausible*. In order to test these types of surprises with complete and homogeneous stories in terms of duration, and to show many of these types of surprises to each participant, 16 narrative television commercials ( $M=40,68$  seconds) were used as stimuli in this study. The experimental design was a 4 groups (3 groups of surprise, and 1 non-surprise group, 4 stories each) fully randomised experiment ( $N=120$ , Age: 18-24). The results showed that surprise had a significant enhancing effect on liking, on the day of the viewing ( $\chi^2(3, N=480)=5.83, p=0.12$ ), and one month after ( $\chi^2(3, N=480)=10.38, p=.016$ ); an ANOVA test showed a significant relation between surprise and the degree of comprehension ( $F(1,480)=12.14, p<.001$ ): stories that elicited surprise were less comprehended ( $M=+2.29, SD=0.50$ ) than the ones that did not surprised ( $M=+2.67, SD=0.89$ ); and non-diegetic surprises elicit better long-term free recall than audiovisual narratives without surprise (50%

recalled,  $AR=3.3$ ). These results suggest a major difference between non-diegetic and diegetic surprises, and they point out that highest levels of surprise elicit a highest degree of liking, even if not fully comprehended. In conclusion, future studies on audiovisual narrative surprises should focus on the relation between the underlying schemata regarding story elements, and the different types of surprising stimuli.

**Keywords**

Emotions, surprise, narrative, diegetic, audiovisual.

## **1. Introduction**

### **1.1. Surprise and audiovisual stimuli**

A wide variety of stimuli have been used for studying surprise and its effects, including written texts (e.g. Ash, 2009), visual and pictorial stimuli (e.g. Niepel, Rudolph, Schützwohl & Meyer, 1994), auditory stimuli (e.g. Niepel *et al.*, 1994), or audiovisual stimuli (e.g. de Wied, Van Boxtel, Posthumus, Goudena & Matthys, 2009; Schaefer, Nils, Sánchez & Philipot, 2010; or Tan & Diteweg, 1996), among others.

It should be stressed the importance of film stimuli to elicit different emotions (for a revision see Schaefer *et al.*, 2010), some of them to elicit surprise, meeting criteria as emotional discreteness, arousal and valence. For example, Gross and Levenson (1995) used to this purpose sequences from the commercial films *Capricorn One* and *Sea of Love*. Sato, Noguchi and Yoshikawa (2007) successfully replicated the experiment with Japanese audiences.

There are also studies that take into account the importance of narrative surprise on feature films (Tan, 1996; Tan & Diteweg, 1996), pointing out that “surprise may be as important for viewer interest as suspense is” (Tan & Diteweg, 1996, p. 175). However suspense has received an experimental larger attention (Vorderer, Wulff & Friedrichsen, 1996), as well as other emotions such as fear (Hoffner & Levine, 2005) or enjoyment of sex and violence (Welsh & Brantford, 2009).

Narrative television commercials include several different structures (Bermejo, 2006): informative structure, persuasive structure, and dramatic structure. Its persuasive power is based on its capacity “to connect with the previous narrative world of the subject through the exhibited story. When television commercials are fictional stories –where *mise-en-scene* of the product, and the audience world are attached to the real world–, audiovisual narratives make possible the encounter of both worlds and an exchange between the two of them” (Bermejo, 2006, p. 107).

Thus, television commercials typically activate two main schemes: 1) Story schema, which comprises character, action, a space-time world schemes; and 2) Advertiser schema, which comprises product, brand, and television commercial schemes.

Narrative television commercials and surprise have been studied related to its advertising efficacy, though it should be noted that besides using the research category of surprise (Yang *et al.*, 2015; Bermejo Berros & López Díez, 2013), they have also been studied using research categories close to it, as incongruency (Törn & Dahlén, 2008), or creativity (Stone *et al.*, 2000).

### **1.2. Narrative surprise**

Some of the few studies that have used surprise stimuli have used sequence clips from movies, instead of complete stories (Gross & Levenson, 1995; Sato *et al.*, 2007). This limits their acknowledge as studies of narrative surprises, as surprises in a narrative discourse affects the whole story, and they are affected by the story as a whole. This suggests the need of studying surprise in complete audiovisual stories, which is one of the purposes of this research. Moreover, the emotions elicited by a story is the result of the interaction of the plot of the story as a whole, and not just by a part of it, it has been pointed out by Baroni (2007).

Firstly, in this study narrations or stories have been understood as those texts with a story grammar containing, at least, the following minimal structure: a setting, that establishes the protagonist, and the space and time of the story, followed by, at least, an episode with three constituents: a beginning constituent, with, at least, an event; a development constituent, consisting of the protagonist’s reaction to the beginning event; and an ending constituent, including the consequences and outcome of the reaction of the protagonist (Mandler, 1984).

Secondly, narrative surprises has been considered as significant plot events discrepant with a schema elicited by the story up to the presentation of the surprising event, which is consistent with Brewer and Lichtenstein's definition of surprise events in stories (Brewer, 1985; Brewer & Lichtenstein, 1981; 1982) as those containing "critical expository or event information early in the event sequence" – "critical in the sense that it is necessary for the correct interpretation of the event sequence" – that "the author withholds from the beginning of the discourse structure without letting the reader" – the spectator, in this study – "know that something has been withheld" (Brewer, 1985, p. 169).

### 1.3. Types of audiovisual narrative surprises

Our first goal was to postulate a classification of audiovisual narrative surprises. In order to do this, some of the criteria proposed by Andrew Ortony and Derek Partridge to classify surprises have been applied: 1) *deducible* and *non deducible* surprises, criterion that led these authors to distinguish between surprise due to expectation failure and surprisingness, and 2) the *practically immutable* or *practically typical* nature of the proposition affected by the surprising event (Ortony & Partridge, 1987).

Surprises have been considered deducible or non-deducible regarding the presentation of the surprising event in the diegetic *world* (the one inhabited by the characters of the story) or in the non-diegetic *world*. When the surprise is presented in the diegetic world, it has been considered a deducible surprise, because it leads to expectation failure of previously (actively or passively) present schemata in the mind of the subject. On the other hand, when the surprising event (affecting the story) is presented non-diegetically, it has been considered a non-deducible surprise because it leads to surprisingness, due to the irruption of a new schema.

Thus, taking these distinctions into account, the terminology suggested by Ortony and Partridge has been adapted to the present study as non-diegetic (non-deducible) and diegetic (deducible) surprises.

Diegetic surprises are the most common type of surprises in audiovisual narratives of any kind, e. g., in the movie *Psycho* the appearance of the mother skeleton is a deducible surprise because it occurs inside the diegetic world.

An example of non-diegetic surprise can be found in the black frame representing the black out, most likely death, of the protagonist in the last episode of the television series *The Sopranos*, David Chase's "Made in America" (Chase, 2007), because a black frame may be a good metaphor for absence of consciousness but it is not an image of the character's mind and, so, it does not belong to the diegetic world.

As for the second criterion suggested by Ortony and Partridge, applied only to diegetic surprises, it has been deemed that the story surprise affected *practically immutable propositions* when it was discrepant with the physical and/or chemical schema of the diegetic world previously elicited by the story, while they affected *practically typical propositions* when the schema-discrepant input did not affected previous physical and/or chemical schema of the diegetic world. The ones of the first type has been called diegetic implausible surprises and, the second, diegetic plausible surprises.

An example of diegetic *implausible surprise* is found in *The Matrix* (Wachowski & Wachowski, 1999), when Neo discovers the reality of the world he inhabits. While an example of *diegetic plausible surprise* is found in the discovering, in Orson Welles' *Citizen Kane*, that "Rosebud" was the name of a sledge.

Thus, regarding the intensity elicited by surprise in television commercials, we formulated two hypothesis (Cf. 2: H1 and H2).

#### 1.4. *Surprise and liking*

Our hypothesis concerning liking was that audiovisual stories with surprise would obtain a higher rate of liking than stories without surprise. Moreover, this study aimed to analyse how the intensity of surprise would affect liking.

Surprise is a neutral hedonically valenced emotion, followed, after the appraisal and analysis of the surprising event, by a positive or negative hedonically valenced emotion, that explains folk talk of positive and negative surprises. Moreover, surprise enhances the intensity of the following emotion. Thus, joy after surprise is more intense than joy without a previous surprise (Desai, 1939).

Though surprise is hedonically neutral, it has been pointed out that an event incongruent –incongruity implies schema-discrepancy– with a preceding schema may be evaluated as positive or negative depending on the level of incongruity (Mandler, 1982). Slight incongruity of the event leads to positive evaluation after assimilation of the event within the previous schema, while severe incongruity may result in a positive evaluation when an alternate schema explains the event, or an effort of accommodation of the event that may be successful, which may lead to an intense positive or negative affect depending on the context, or unsuccessful, in which case it elicits an intense negative evaluation.

This means that moderate incongruities elicit positive or negative evaluation depending on the context, while severe incongruities may not be solved, which result in an intense negative affect.

On the other hand, congruity elicits positive evaluation (Mandler, 1982). Even if a more positive evaluation is elicited by successfully assimilated or accommodated incongruity (Schellenberg, Peretz & Vieillard, 2008).

In a narrative context, Mandler's hypothesis is consistent with Brewer and Lichtenstein's structural affect theory of stories which predicts that unsolved narrative surprises will reduce liking of stories while solved narrative surprises elicit an increased liking of stories.

While watching feature films, viewer and maker have made a "pragmatic contract" (Tan, 1996, p. 243), which means that "the maker of the traditional feature film is out to entertain the audience, while the viewers are there precisely because they want to be entertained". Thus, when watching a feature film, a negative emotion such as fright, result of surprise and fear, may be, in fact, positively evaluated as interesting and entertaining, which and thus, positively enhancing. Tan has also pointed out the role of surprise in film stories as reinforcement of other emotions, such as suspense.

In the field of advertising, it has been found that advertisements which included an event that was incongruent with the advertisement schema previously activated by the viewer are preferred to those congruent with that schema (Lee & Mason, 1999; Loef & Verlegh, 2002). Besides, viewers prefer creative television commercials to non-creative television commercials (Ang & Low, 2000).

On this matter, we formulated one hypothesis (Cf. 2: H3).

#### 1.5. *Surprise and recall*

Concerning recall, it was expected that audiovisual stories with surprise would achieve a higher recall because surprise has been reported to enhance memory and, specifically, free recall, due to increased processing of schema-discrepant events compared with non schema discrepant events (Niepel *et al.*, 1994; Stiensmeier-Pelster, Martini & Reizenzein, 1995). According to the *cognitive psychoevolutionary model of surprise*, this processing comprises several subprocesses: verification of the schema discrepant, analysis of the causes of the unexpected event, assessment of its relevance for ongoing actions, and, sometimes, assessment of moral significance (Meyer, Reizenzein & Schützwohl, 1997). However this memory advantage of surprise, some studies have also found surprise as related to memory disruption (Hunt, Reed & Worthen, 2006).

Other experiments which have studied concepts regarding the stimuli that overlap with surprise or unexpected events also support enhanced memory performance. Thus, studies on incongruent stimuli (Michelon & Snyder, 2006), atypical stimuli (Mandler, 1984), and secondary distinctiveness stimuli and bizarre stimuli (Hunt *et al.*, 2006). For some of these concepts it is still controversial to what extent the effect of surprise is responsible for, at least, part of the memory performance advantage (Hirshman, Whelley & Palij, 1989, Worthen, 2006).

Another factor that affects recall of surprising events is the intensity of the experience of surprise. The more intense the surprise experienced by a subject is, the more likely the subject is going to recall the surprising event (Thorson & Friestad, 1989). As a general statement, it can be said that “the more a value of a variable is expected (that is, the more typical or higher the probability of occurrence of that value) the less well it is encoded or remembered” (Mandler, 1984, p. 103). Even to the extent that not understanding the semantic relation between two items may enhance memory performance (Sokolov, 1963, cited by Hirshman *et al.*, 1989).

It is also has to be taken on account the retention interval, because it may change the memory advantage for unexpected events, as it is observed for the effect of typicality on memory. According to Smith and Graesser (1981), “as the retention interval increases, the tagged atypical actions in the trace become less accessible because retrieval becomes more dependent upon the generic schema” (p. 557). Thus, the increased recall for atypical events observed for short-term retention intervals may be reduced or neutralized by a long-term retention interval.

Regarding advertising, Heckler and Childers (1992) found that advertisements that used unexpected and relevant information were more recalled than those containing unexpected and irrelevant information, or expected information (relevant or irrelevant). And Törn and Dahlén (2008) found that brand-incongruent television commercials increased brand recall.

Thus, regarding surprise and recall, we formulated one hypothesis (Cf. 2: H4).

### **1.6. Surprise and comprehension**

Regarding comprehension, it was aimed to study how comprehension of narrative surprises affected recall and liking. Comprehension is a cognitive dimension related to both recall and liking. On one hand, recall may be enhanced by not resolving a surprise or being unsuccessful in accommodating an unexpected event because it leads to deeper and longer processing of the event (Sokolov, 1963) as well as resulting in longer lasting subsequent emotions like confusion or boredom (D’Mello & Graesser, 2011, p. 1301): “Surprise that occurs in response to novelty and delight when an intermediate learning goal is achieved is expected to be quite brief [...]. In contrast, confusion and frustration occur when the discrepancy or novelty triggers an impasse that blocks an important superordinate learning goal (e.g., solving a difficult problem or understanding a complex topic).”

On the other hand, successfully accommodation of severe incongruent events elicit an intense negative evaluation of the event (Mandler, 1982). Likewise, in the context of narrative texts, according to the structural-affect theory of stories, the resolution of story surprises lead to increased story liking and enjoyment (Brewer, 1984).

On this matter we formulated a research question (Cf. 2: RQ1)

## **2. Method**

The main objective of this research is to test the general hypothesis that the two variables applied to the audiovisual narrative surprise stimuli, namely, diegetic/non-diegetic and implausible/plausible, elicit significant different levels of surprise intensity, recall, liking and comprehension when applied to whole stories.

Thus, we have formulated the following hypothesis, and one research question:

H1: Narrative television commercials with a non-diegetic surprise will elicit a higher intensity of surprise than narrative television commercials with a diegetic surprise.

H2: Narrative television commercials with a diegetic implausible surprise will elicit a higher intensity of surprise than narrative television commercials with a diegetic plausible surprise.

H3: The subjective experience of surprise enhances preference of a television commercial.

H4: Narrative television commercials with surprise will obtain higher long-term recall values than narrative television commercials with no surprise.

RQ1: how does comprehension of narrative surprises in television commercials affect recall and liking?

### 2.1. Participants

Participants were 120 undergraduate students (mean age: 19 years and 11 months; range: 18–24 years), volunteers, from the grade in Advertising of the University of Valladolid.

### 2.2. Materials

The selection of narrative audiovisual stimuli was made applying two criteria: 1) they should be whole stories, and not just clips from a whole story; and 2) both the type of schema and the type of surprise proposed should be able to be controlled. Besides, it was taken into account if the stories selected were appropriate to show several stimuli of each kind of surprise to each participant in the same research session. Thus, there were used television commercials with brief stories, which met both criteria.

The stimuli showed to the participants were 16 narrative television commercials, in Flash Video format (.flv). The duration of the videos was 20 to 60 seconds ( $M=40.68$  s), chosen from a video archive of around 2000 television commercials, the historical video archive of the Latin-American Festival of Advertising “El Sol”. All of the commercials were in Spanish and had been broadcasted on television several years before (between 2006 and 2016), and thus, they were unknown by the subjects.

There were four groups of stories: A) non-diegetic surprises; B) diegetic implausible surprises; C) diegetic plausible surprises; and D) non-surprise group. Each group comprised four stories.

An initial selection of the stories was made by the researchers and the final choice was made in accordance with external narrative experts attending to the criteria of narrativity and type of surprise.

The titles of the commercials (in the historical video archive of Festival “El Sol”) were: A) “MTV–Punk Not Dead”; “TVE– Spain–Serbia football match Kitchen”; “Impulse– Wedding Ring”; “Monopoly–Prison”; B) “Airtel–Myrrh”; “Fritos Lay– China”; “TVE Grapes New Year’s Day”; “Ikea–Suicidal Sofa”; C) “Cyloop– Horse”; “El País–Prison”; “Campofrío– Dog”; “TVE– Christmas Lotto”; D) “La Lechera– Recipes”; “Audi– Curves”; “Gala Loewe–Dress for a gala”; “Wipp Turbo– Ester and friend”.

### 2.3. Design

The experimental design was a four groups (A, B, C, D) fully randomised study. The videos were randomly presented using the stimulus presentation software *Superlab 4*. The dependent variables were: subjective experience of surprise (after watching the video); intensity of the subjective experience of surprise (after watching the video); free recall (one month later); liking (after watching the video) and comprehension (after watching the video). They were all measured by self-report from the participants and can be classified in two types:

**Qualitative variables:** 1) subjective experience of surprise (nominal dichotomous variable, yes/no). Participants were not directly asked if they had felt surprise but what emotions they had felt. Subjects who had not mentioned feeling surprise after watching a particular story were asked after watching all stories if they had felt surprise while watching that specific story; 2) liking, measured asking a) after watching all stories, which were the most liked and most disliked story, and b) one month after watching the stories, asking which was the most liked story.

**Quantitative (scale) variables:** 1) intensity of the experience of surprise, measured on a 1-7 point scale (1, minimum; 7, maximum); 2) free recall, measured asking what commercials they spontaneously recalled; 3) liking, measured immediately after watching each story on a -5 to +5 scale, including 0 (-5, extreme disliking, +5, extreme liking); 4) comprehension, measured on a 0-3 scale (0, not at all; 3, totally).

## 2.4. Procedure

On account of the complexity of the studied variables, we run a pre-test to control the knowledge and relation of the participants with the advertised brand before the viewing and the test.

Participants were shown the videos singly in a laboratory room at University of Valladolid. The researcher was always present. First, participants were told that this experiment wanted to test the comprehension of television commercials, to avoid making explicit the purpose of studying surprise. They were instructed about the procedure of the experiment. The researchers run the stimulus presentation software *Superlab 4*, which randomly presented the stimulus. Participants watched the 16 videos on a 21" computer screen. After watching each video, participants were asked about the emotions and its intensity, the comprehension, and how much they liked the video. After watching the 16 videos, they were asked about surprise and its intensity if it has not been mentioned, and about the most liked and most disliked story. The experiment lasted around one hour for each participant. One month later, participants were contacted on the phone and asked what stories they recalled and which was the story they had liked the most.

## 3. Results

Several statistical tests were used to analyse the data: 1) Pearson chi-squared test ( $\chi^2$ ) and contingency tables to analyse the relation between two nominal variables; 2) analysis of the variance (ANOVA) to study the relation between a nominal variable and a scale variable; 3) and correlation to study the relation between two scale variables.

### 3.1. Intensity of surprise

A first analysis tested whether the stories belonging to the three surprise groups had effectively elicited surprise significantly higher than the stories from the non-surprise group.

There was a significant larger number of participants that experienced surprise watching stories of surprise groups, compared to stories of non-surprise group,  $\chi^2(3, N=480)=382.46$ ,  $p<.001$ . All of non-diegetic surprise stories (100%), almost all (95%) of diegetic implausible surprise stories, and almost all (94 %) of the diegetic plausible surprise group, elicited surprise, while few of the non-surprise group (6%) elicited it.

Similarly, an ANOVA test showed that stories from surprise groups elicited a higher intensity of surprise than stories from the non-surprise group,  $F(3,480)=273.45$ ,  $p<.001$ . Stories from non-diegetic surprise group elicited the highest intensity of surprise ( $M=5.49$ ,  $SD=1.43$ ), followed by diegetic implausible surprises ( $M=4.88$ ,  $SD=1.89$ ), diegetic plausible surprises ( $M=4.65$ ,  $SD=1.96$ ) and stories from the non surprise group ( $M=0.21$ ,  $SD=0.96$ ).

Consequently, according to these results, H1 was confirmed: the distinction between diegetic and non-diegetic surprise was proved significant. Regarding H2, results are more

complex, as results partially confirmed our expectations, as we will be discuss in section 4 (Discussion).

### 3.2. Liking

#### 3.2.1. Degree of liking

The relation between the elicitation of surprise and liking was proved significant by an ANOVA test,  $F(1,480)=58.35$ ,  $p<.001$ . Stories which elicited surprise were more liked ( $M=+2.64$ ,  $SD=2$ ) than those which didn't elicited surprise ( $M=+1.07$ ,  $SD=1.89$ ). Similarly, a Pearson correlation proved that intensity of surprise significantly correlated with liking,  $r(480)=.41$ ,  $p<.001$ .

Likewise, surprise had a significant enhancing effect on most liked and disliked stories. Consequently, H3 was also confirmed.

#### 3.2.2. Most liked/disliked story

**Most liked story on the day of the viewing:** On the day of the viewing, stories that elicited surprise were more liked than stories that didn't elicited it,  $\chi^2(1, N=480)=4.36$ ,  $p=.037$ . A higher percentage (8%,  $AR=2.1$ ) of all stories that elicited surprise was chosen as most liked story compared to same percentage for stories that hadn't elicited surprise (2%,  $AR=-2.1$ ).

An ANOVA test showed a significant relation between intensity of surprise and most liked stories the day of the viewing,  $F(1,480)=8.19$ ,  $p=.004$ . Most liked stories had significantly elicited a higher level of surprise ( $M=5.13$ ,  $SD=2.34$ ) than stories not chosen as most liked,  $M=3.72$ ,  $SD=2.64$ .

**Most liked story one month after the viewing:** One month after the viewing, the relation between surprise and liking was also significant,  $\chi^2(1, N=480)=6.34$ ,  $p=.012$ . A higher percentage (8%,  $AR=2.5$ ) of all stories that elicited surprise were chosen as most liked stories compared with same percentage for stories that had not elicited surprise (2%,  $AR=-2.5$ ). Similarly, an ANOVA test revealed intensity of surprise elicited by stories chosen as most liked ( $M=4.93$ ,  $SD=2.24$ ) was significantly higher ( $F(1,480)=5.89$ ,  $p=.016$ ) than intensity elicited by stories not chosen as most liked stories ( $M=3.73$ ,  $SD=2.65$ ).

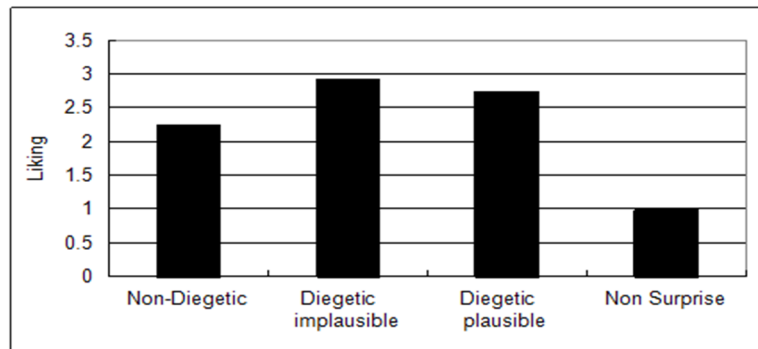
**Most disliked story on the day of the viewing:** The relation between experience of surprise and most disliked stories was also significant,  $\chi^2(1, N=480)=4.82$ ,  $p=.028$ . The percentage of most disliked stories that didn't elicited an experience of surprise was more than approximately double (10%,  $AR=2.2$ ) than the percentage of most disliked stories that did elicited surprise (5%,  $AR=-2.2$ ). As for the relation between most disliked stories and intensity of surprise, again a significant level,  $F(1,480)=2.75$ ,  $p=.098$ , stories chosen as most disliked stories elicited lower intensity of surprise ( $M=3.03$ ,  $SD=2.86$ ) than stories not chosen as most disliked stories ( $M=3.86$ ,  $SD=2.62$ ).

#### 3.2.3. Types of surprise and liking

**Degree of liking:** An ANOVA test showed that stories belonging to surprise groups elicited significantly higher degrees of liking than stories of the non surprise group,  $F(3,480)=23.83$ ,  $p<.001$ ). On a -5 to +5 point-scale (including 0) diegetic implausible surprises elicited the highest degree of liking ( $M=+2.93$ ,  $SD=1.59$ ), followed by diegetic plausible surprises ( $M=+2.74$ ,  $SD=1.62$ ), non-diegetic surprises ( $M=+2.25$ ,  $SD=2.63$ ) and non surprise group stories ( $M=+0.99$ ,  $SD=1.83$ ).



**Figure 1:** Liking and groups of surprise/non surprise.



Note: Liking was measured on a scale from +5 to -5.

**Table 1:** Liking and groups of surprise/non surprise.

Surprise/Non surprise groups	Liking		
	n	M	SD
Non diegetic surprise	120	2.25	2.63
Diegetic implausible surprise	120	2.93	1.59
Diegetic plausible surprise	120	2.74	1.62
Non surprise	120	0.99	1.83

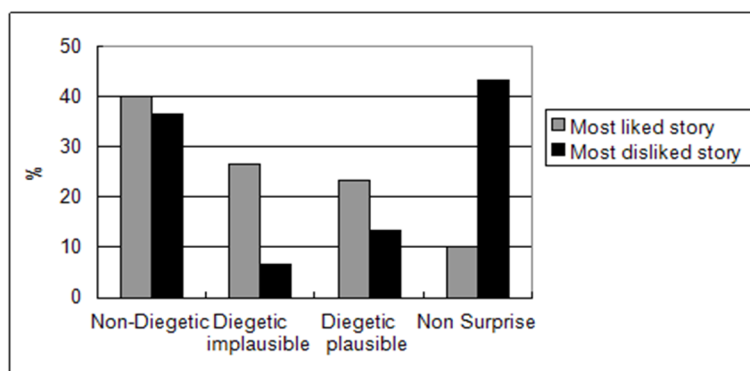
Note: Liking was measured on a scale from +5 to -5.

**Most liked story:** Story belonging to the three types of surprise were significantly more chosen as most liked stories both on the day of the viewing ( $\chi^2(3, N=480)=5.83, p=0.12$ ) (See Table 4) and one month after,  $\chi^2(3, N=480)=10.38, p=.016$ ) (See Figure 2)

On the day of the viewing, 40% of participants chose a non-diegetic surprise story as most liked story followed by diegetic implausible surprises (26.6%), diegetic plausible (23.3%) and non-surprise stories (10%). One month after, non-diegetic was also the highest (43.3%), followed by both diegetic surprise stories (26.6%, both types) and non-surprise stories (3.3%).

**Most disliked story:** As for most disliked story, the relation was also significant ( $\chi^2(3, N=480)=12.09, p=.007$ ). The highest number of most disliked story belonged to the non-surprise group (43.3%), followed by non-diegetic surprise group ( $N=36.6%$ ), diegetic plausible surprises ( $N=13.3%$ ), and diegetic-implausible surprises ( $N=6.6%$ ).

**Figure 2:** Most liked/disliked story on the day of the viewing.



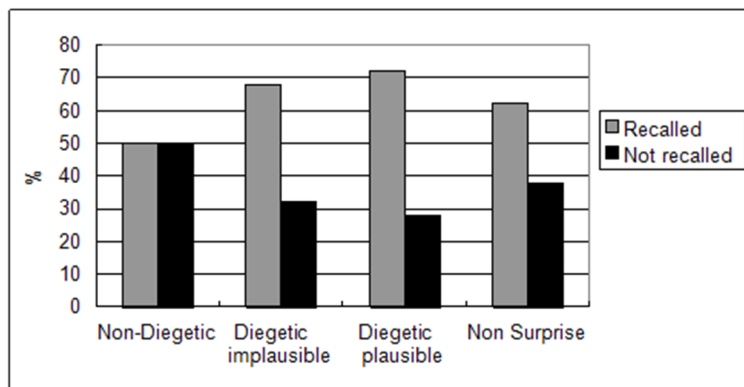
Note: For most liked story:  $\chi^2(3, N=480)=5.83, p=0.12$ . For most disliked story:  $\chi^2(3, N=480)=12.09, p=.007$ .

### 3.3. Recall

Results showed that, analysing both surprise and non-surprise groups, the elicitation of surprise did not have a significant influence on long-term recall,  $\chi^2(1, N=480)=0.04$ ,  $p=0.832$ ). Neither had the intensity of surprise on recall,  $F(1,480)=0.10$ ,  $p=0.757$ ). A similar intensity of surprise on a 1-7 point-scale was elicited by recalled ( $M=3.78$ ,  $SD=2.62$ ) and non recalled stories ( $M=3.85$ ,  $SD=2.682$ ). This means that H4 was not confirmed.

When only stories from surprise groups were analysed, the relation between types of surprise and long-term recall was proved significant ( $\chi^2(3, N=480)=13.64$ ,  $p=0.003$ ), although only for the group of non diegetic surprises (50% recalled,  $AR=3.3$ ) and for the group diegetic plausible surprises (recalled: 28%,  $AR=-2.3$ ), while it was not significant for the diegetic implausible group (recalled: 32%,  $AR=-0.3$ ) and for the non-surprise group (recalled: 38%,  $AR=0.3$ ). (See Figure 3).

**Figure 3:** Long-term recall (One month after the viewing) and groups of surprise/non surprise.



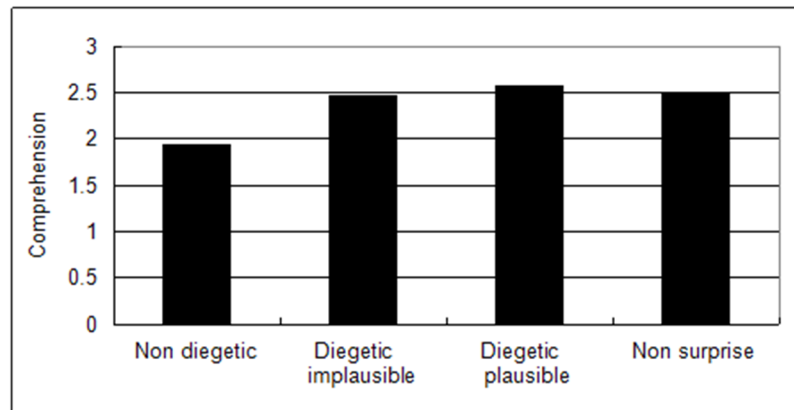
Note.  $\chi^2(3, N=480)=13.6$ ,  $p<0.003$ .

### 3.4. Comprehension

An ANOVA test showed a significant relation between surprise and the degree of comprehension ( $F(1,480)=12.14$ ,  $p<.001$ ). On a 0 to 3 point scale (where 0 is “I did not understand it at all,” and 3 “I understood it completely”), stories that elicited surprise obtained a lower level of comprehension ( $M=+2.29$ ,  $SD=0.50$ ) than stories that did not elicited surprise ( $M=+2.67$ ,  $SD=0.89$ ).

Groups of surprise and degree of comprehension were also significantly related,  $F(3,480)=18.36$ ,  $p<.001$ . Non-diegetic surprises obtained the lowest level of comprehension ( $M=1.93$ ,  $SD=1.02$ ), while the rest of the groups rated similar comprehension levels: diegetic-implausible surprises ( $M=2.46$ ,  $SD=0.60$ ), diegetic plausible surprises ( $M=2.58$ ,  $SD=0.59$ ), and non-surprise group ( $M=2.48$ ,  $SD=0.673$ ).

**Figure 4:** Comprehension and groups of surprise/non surprise.



Note. Comprehension was measured on a scale from 0 to 3.

**Table 2:** Comprehension and groups of surprise/non surprise.

	Comprehension*				
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>CI 95%</i>	
<b>Groups of surprise/ non surprise</b>				LL	UL
<i>Non diegetic</i>	120	1.93	1.019	1.75	2.12
<i>Diegetic implausible</i>	120	2.46	0.607	2.35	2.57
<i>Diegetic plausible</i>	120	2.58	0.589	2.47	2.68
<i>Non surprise</i>	120	2.48	0.673	2.36	2.61

Note. Comprehension was measured on a scale from 0 to 3. N: participants; CI 95%: Confidence interval. LL: Lower limit. UL: Upper limit.

## 4. Discussion

### 4.1. Diegetic and non-diegetic surprises

First of all, the distinction between surprise and non-surprise groups proved effective to elicit significantly differentiated intensities of the feeling surprise. This supports the validity of the selection of stories with the presence of an event, important for the comprehension of the story, schema-discrepant with a preceding activated schema to elicit the emotion of surprise meeting a criterion of discreteness.

This result was not consistent with Ortony and Partdrige classification of surprises and their predictable intensities of surprise elicited, as they stated that both deducible and non deducible surprises could obtain the highest level of intensity (Ortony & Partridge, 1987). But it did was consistent with Maguire and Maguire (2009) consideration on the different intensities of surprises depending on the difficulty to integrate an unexpected event into a previous or new schema. The intensity of surprise is higher when the difficulty of integrating is higher. In the case of non-diegetic surprises, this difficulty is higher, as it is a different level of story schema (non-diegetic vs. diegetic).

It is also consistent with Lorini and Castlefranchi's (2009) major distinction between mismatch expectations surprises and astonishment, who predicted a higher intensity of surprise for the latter, which have more in common with diegetic that with non-diegetic surprises.

Regarding H2, the hypothesis was partially confirmed in the sense that diegetic implausible surprises did elicit a higher intensity of surprise than diegetic plausible surprises,

but the difference was not significant. An explanation to this result may be the duration of the stories presented to the participants: 20 to 60 seconds. Diegetic implausible surprises were based on the discrepancy between the world schema activated in the beginning of the story and the sudden irruption of a new world schema, discrepant with the physical and/or chemical schema of the diegetic world previously elicited by the story. Thus, as the time for building the first world schema was very short, the activated expectations by the viewer may have been weaker than in longer stories. And, consequently, the intensity of the surprise would be weaker. Further research should be done looking for key aspects of diegetic surprises that may elicit significant surprise intensity variations, and time for building world schemes may be a significant one.

Thus, a major distinction of audiovisual surprises may be made between surprises affecting the story presented at the diegetic world and those presented at the non-diegetic world. It should be stressed that this distinction is different than the one proposed by Tan between F-emotions and A-Emotions (Tan, 1996), because the latter do not affect the plot of the story. For instance, the surprise felt by audiences when watching for the first time a 3-D movie, which may be categorised as an A-emotion, does not affect the plot nor its comprehension. While the non-diegetic audiovisual narrative surprises proposed in this study *do* affect the plot.

#### 4.2. *Not enhanced recall*

According to the findings of this research, H4 was not confirmed. That is, stories that elicited surprise or high intensities of surprise were not more recalled one month after the exposure to the stimuli. This may suggest that there is a significant influence of retention interval pointed out by Smith and Graesser (1981), a consideration supported by the fact that the non-surprise groups stories not only did not included important schema discrepant events, but also were selected because of the typicality, and this typicality may have played a significant role in the retrieval process. In this sense, it would be useful to study more thoroughly the effect on long-term and short-term recall of expected and unexpected audiovisual story events.

Nevertheless, when analysing different types of surprise groups, it was observed that non-diegetic surprise stories were significantly better recalled ( $AR=3.3$ ), as half of the group stories ( $N=60$ ) were recalled, compared with the rest of the groups, where the number of stories recalled was always lower than the not recalled stories (See Figure 3).

This finding is consistent with the enhancing effect on memory of a low level of comprehension due to a more intense processing of the stimulus (see *supra*), as non diegetic surprise stories showed the lower level of comprehension of all groups. This is possibly related to the fact that they also showed the highest levels of intensity of surprise, which is consistent with the findings of other authors (Mandler, 1984; Thorson & Friestad, 1989).

Another explanation for better recalling non-diegetic surprise stories is that when the participant is watching it, there is a search for comprehension and integration of the discrepant event in the activated story schema, Thus there is a more thorough cognitive processing which leads to a higher recall.

#### 4.3. *The paradox of non-diegetic surprises*

The results supported studies that point out the more positive evaluation of stories with surprise, as all of the measurements of liking in this study showed a significant effect of the elicitation and the intensity of surprise.

It is remarkable that non-diegetic surprise stories, which elicited the highest intensity of surprise, were *the* most liked stories and the second most disliked stories. This was probably due to the fact that they were also the least understood stories, or, in other words, the stories that elicited more confusion, which has a negative influence on liking.

Subjects need to make stories congruent and to integrate story events into a major story schema. In the case of non-diegetic surprises, the subject attempts a congruent representation of its meaning. The higher the surprise, the more intense is the attempt. When congruency is achieved, positive evaluation results. When congruency is not achieved, the process elicits a response of confusion, related to a negative emotional evaluation. Besides, when the subject is not able to solve the incongruity, he has a feeling of personal failure and incompetence. This may lead to the tendency to make an external attribution, tagging the story as a bad story or just assuming that he dislikes it.

Likewise, this may support Mandler's hypothesis that severe incongruities elicit more intense evaluations, both positive, if successfully accommodated, and negative, if not successfully (Mandler, 1982). As well as for the effect of not resolved surprises on decreasing liking (Brewer & Lichtenstein, 1981; 1982).

As for the negative evaluations of non-surprise stories, which were the most disliked and the least liked stories, it may be consistent with a negative effect of over-familiarity or boredom (Schellenberg *et al.*, 2008).

Finally, these results about non-diegetic surprises can also be related to processes of arousal enhancement due to confronting conflictive surprising events (Berlyne, 1960), and recent research on psychology of interest (Silvia, 2006; Renninger & Hidi, 2011).

We consider that this research provides a significant understanding of surprise-based narratives in television commercials. Nevertheless, some limitations should be noted. First, our research participants were students of advertising, a population who has a deeper knowledge of advertising, television commercials and their narratives than general audiences. Future research may be able to apply our results to general population. Second, ad exposure was intentional rather than incidental, as participants watched the commercials within an experimental environment, not in between other television programs, or media content. Future research may examine the insights provided by our research in real life contexts. Third, as actual television commercials were used, extraneous variations of the independent variable cannot be excluded. Future research may use more controlled stimuli to avoid this risk.

## **5. Conclusions and Future Prospects**

Surprises are one of the most important emotions elicited and expected in audiovisual stories of all kinds (feature films, videogames, narrative television commercials, etc.). Though cognitive and experimental approaches to audiovisual narrative surprises are being made, they are still scarce, especially when compared with research on another emotion present in audiovisual stories, namely, suspense. To our knowledge, the present study is the first to analyse different types of surprise based on psychological criteria in audiovisual stories and, thus, it represents an advance in this field.

Firstly, the findings of this study suggest a major difference between two types of the surprise, non-diegetic and diegetic surprises, with predictable differentiated levels of intensity in the feeling of surprise. Secondly, it confirms positive effect on liking for moderate levels of surprise intensity but, for high intensities of surprise, as for non-diegetic surprises, it may lead to negative evaluations of the story, probably because of eliciting incomprehension or confusion. Paradoxically, the highest levels of surprise intensity also elicit the most positive story evaluations. This ambiguity makes of non-diegetic surprises a double-edge narrative strategy that, carefully used achieves the highest rates of liking while, miscalculated, provokes the lowest rates of liking.

Future research should study more deeply the mechanisms that lead to these different types of stimuli assessing, and analyse their relation with the nature of involved schemata, as well as the influence of emotions elicited by different story elements such as characters, gender, and spatial and temporal schemata.

## References

- Ang, S. H. & Low, S. Y. M. (2000). Exploring the Dimensions of Ad Creativity. *Psychology and Marketing*, 17 (10), 835-854.
- Ash, I. K. (2009). Surprise, memory, and retrospective judgment making: Testing cognitive reconstruction theories of the hindsight bias effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 916-933. <https://www.doi.org/10.1037/a0015504>
- Baroni, R. (2007). *La tension narrative*. Paris: Seuil.
- Berlyne, D. E. (1960): *Conflict, Arousal, and Curiosity*. New York: McGraw Hill.
- Bermejo-Berros, J. (2006). La eficacia de la publicidad narrativa. Nuevos caminos para el anunciante. *Publifilia*, 9, 93-107.
- Bermejo-Berros, J. & López Díez, J. (2013). Sorpresa y diégesis en la narrativa audiovisual. In P. Gómez Martínez (Ed.), *Teorías y aplicaciones narrativas* (pp. 131-160). Madrid: Icono 14.
- Brewer, W. F. (1985). The story-schema: Universal and culture-specific properties. In D. R. Olson, N. Torrance & A. Hildyard (Eds.), *Literacy, language and learning: The nature and consequences of reading and writing* (pp. 167-194). Cambridge, UK: Cambridge University Press.
- Brewer, W. F. & Lichtenstein, E. H. (1981). Event schemas, story schemas, and story grammars. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX*. (pp. 363-379). Hillsdale, NJ: Lawrence Erlbaum.
- Brewer, W. F. & Lichtenstein, E. H. (1982). Stories are to entertain: A structural-affect theory of stories. *Journal of Pragmatics*, 6, 473-486. [https://www.doi.org/10.1016/0378-2166\(82\)90021-2](https://www.doi.org/10.1016/0378-2166(82)90021-2)
- Desai, M. M. (1939). Surprise: a historical and experimental study. *British Journal of Psychology, Monograph Supplements*, 22, 1-124.
- De Wied, M., Van Boxtel, A., Posthumus, J. A., Goudena, P. P. & Matthys, W. (2009). Facial EMG and heart rate responses to emotion-inducing film clips in boys with disruptive behavior disorders. *Psychophysiology*, 46(5), 996-1004. <https://www.doi.org/10.1111/j.1469-8986.2009.00851.x>
- D'Mello, S. & Graesser, A. (2011). The half-life of cognitive-affective states during complex learning. *Cognition & Emotion*, 25(7), 1299-1308. <https://www.doi.org/10.1080/02699931.2011.613668>
- Gross, J. J. & Levenson, R. W. (1995). Emotion elicitation using film. *Cognition and Emotion*, 9(1), 87-108. <https://www.doi.org/10.1080/02699939508408966>
- Heckler, S. E. & Childers, T. L. (1992). The role of expectancy and relevancy in memory for verbal and visual information: What is incongruency? *The Journal of Consumer Research*, 18(4), 475-492.
- Hirshman, E., Whelley, M. & Palij, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory and Language*, 28, 594-609. [https://www.doi.org/10.1016/0749-596X\(89\)90015-6](https://www.doi.org/10.1016/0749-596X(89)90015-6)
- Hoffner, C. A. & Levine, K. J. (2005). Enjoyment of mediated fright and violence: A meta-analysis. *Media Psychology*, 7, 207-237. [https://doi.org/10.1207/S1532785XMEP0702\\_5](https://doi.org/10.1207/S1532785XMEP0702_5)
- Hunt, R. & Worthen, J. B. (Eds.) (2006). *Distinctiveness and Memory*. Oxford, UK: University Press. <https://www.doi.org/10.1093/acprof:oso/9780195169669.001.0001>
- Lee, Y. H. & Mason, C. (1999). Responses to information incongruency in advertising: the role of expectancy, relevancy, and humor. *Journal of Consumer Research*, 26(2), 156-169.
- Loef, J. & Verlegh, P. W. J. (2002). Cognitive and affective consequences of two types of incongruent advertising. *ERS-2002-42-MKT*, Erasmus Research Institute of Management (ERIM).

- Lorini, E. & Castelfranchi, C. (2006). The unexpected aspects of surprise. *International Journal of Pattern Recognition and Artificial Intelligence*, 20(6), 817-835. <https://www.doi.org/10.1142/S0218001406004983>
- Maguire, P. & Maguire, R. (2009). Investigating the difference between surprise and probability judgments. *Proceedings of the Thirty-First Conference of the Cognitive Science Society*, Amsterdam (pp. 2359-2364). Mahwah, NJ: Erlbaum.
- Mandler, G. (1982). The structure of value: Accounting for taste. In H. M. S. Clarke & S. T. Fiske (Eds.), *Affect and Cognition: The 17<sup>th</sup> Annual Carnegie Symposium on Cognition* (pp. 3-36). Hillsdale, NJ: Lawrence Erlbaum.
- Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. Hillsdale, NJ: Lawrence Erlbaum.
- Meyer, W.-U., Reisenzein, R. & Schützwohl, A. (1997). Towards a process analysis of emotions: The case of surprise. *Motivation and Emotion*, 21, 251-274. <https://www.doi.org/10.1023/A:1024422330338>
- Michelon, P. & Snyder, A. Z. (2006). Neural correlates of incongruity. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 361-380). New York, NY: Oxford University Press. <https://www.doi.org/10.1093/acprof:oso/9780195169669.001.0001>
- Niepel, M., Rudolph, U., Schützwohl, A. & Meyer, W.-U. (1994). Temporal characteristics of the surprise reactions induced by schema-discrepant visual and auditory events. *Cognition & Emotion*, 8(5), 433-452. <https://www.doi.org/10.1080/02699939408408951>
- Ortony, A. & Partridge, D. (1987). Surprisingness and expectation failure: What's the difference? *IJCAI 87 Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, 1, 106-108.
- Reisenzein, R. & Meyer, W.-U. (2009). Surprise. In D. Sander & K. R. Scherer (Eds.), *Oxford companion to the affective sciences* (pp. 386-387). Oxford: Oxford University Press.
- Renninger, K. A. & Hidi, S. (2011) Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, 46(3), 168-184. <https://www.doi.org/10.1080/00461520.2011.587723>
- Schaefer, A., Nils, F., Sánchez, X. & Philipot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion*, 24(7), 1153-1172. <https://www.doi.org/10.1080/02699930903274322>
- Sato, W., Noguchi, M. & Yoshikawa, S. (2007). Emotion elicitation effect of films in a Japanese sample. *Social Behavior and Personality*, 35(7), 863-874. <https://www.doi.org/10.2224/sbp.2007.35.7.863>
- Schellenberg, E. G., Peretz, I. & Viellard, S. (2008). Liking for happy –and sad– sounding music: Effects of exposure. *Cognition and Emotion*, 22, 218-237. <https://www.doi.org/10.1080/02699930701350753>
- Silvia, P. J. (2006). *Exploring the psychology of interest*. New York, NY: Oxford University Press.
- Smith, D. A. & Graesser, A. C. (1981). Memory for actions in scripted activities as a function of typicality, retention interval, and retrieval task. *Memory and Cognition*, 9(6), 550-559.
- Sokolov, E. N. (1963). *Perception and the conditioned reflex*. New York: Pergamon.
- Stiensmeier-Pelster, J., Martini A. & Reisenzein, R. (1995). The role of surprise in the attribution process. *Cognition and Emotion*, 9(1), 5-31. <https://www.doi.org/10.1080/02699939508408963>
- Stone, G., Besser, D. & Lewis, L. E. (2000). Recall, Liking, and Creativity in TVCommercials: A New Approach. *Journal of Advertising Research*, 40(3), 7-18.
- Tan, E. S. (1996). *Emotion and the structure of narrative film*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Tan, E. S. & Ditleweg, G. (1996). Suspense, Predictive Inference, and Emotion in Film Viewing. In P. Vorderer, H. J. Wulff & M. Friedrichsen (Eds.), *Suspense* (pp. 149-188) Mahwah, NJ: Erlbaum.

- Thorson, E. & M. Friestad (1989). The Effects of Emotion on Episodic Memory for Television Commercials. In P. Cafferata & A. Tybout (Eds.), *Cognitive and Affective Reactions to Advertising* (pp. 305-325). Lexington, MA: Lexington Books.
- Törn, F. & Dahlén, M. (2008). Effects of Brand Incongruent Advertising in Competitive Settings. *European Advances in Consumer Research*, 8, 234-39.
- Welsh, A. & Brantford, L. (2009). Sex and violence in the slashed horror film: A content analysis of gender differences in the depiction of violence. *Journal of Criminal Justice and Popular Culture*, 16, 1-25.
- Worthen, J. B. (2006). Resolution of discrepant memory: An explanation of the effects of bizarreness on memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and Memory* (pp. 133-156). New York: Oxford University Press.  
<https://www.doi.org/10.1093/acprof:oso/9780195169669.001.0001>
- Yang, T., Lee, D.-Y., Kwak, Y., Choi, J., Kim, C. & Kim, S.-P. (2015). Evaluation of TV commercials using neurophysiological responses. *Journal of Physiological Anthropology*, 34:19. <http://www.doi.org/10.1186/s40101-015-0056-4>