



Everybody's a Critic; Influencability of motion picture evaluations from expert and consumer reviews

Master thesis by:

Ruud Jacobs

0160970

MPS-HFM

University of Twente

Coordinators:

Dr. A. Heuvelman, GW-MCO

Dr. S. Ben Allouch, GW-MCO

August 29, 2012

Everybody's a Critic; Influencability of motion picture evaluations from expert and consumer reviews

Master's thesis by Ruud S. Jacobs, 2012

Coordinators: Ard Heuvelman & Somaya Ben Allouch

Samenvatting

Recensies van entertainmentproducten worden voornamelijk in drie velden wetenschappelijk onderzocht; marketingonderzoek bekijkt effecten van recensies op winstgevendheid van bijvoorbeeld films, terwijl communicatiewetenschappers zich vooral richten op de nieuwe trend van online consumentenrecensies. Een kleiner, derde onderzoeksveld richt zich op de invloeden van recensies op de filmwaardering van hun publiek. Het huidige onderzoek combineerde deze gebieden in een exploratief model dat de invloeden vergeleek van professionele critici en gebruikersrecensies op de entertainmentervaring van een consument, waarbij diens involvement werd gezien als moderator. Een experiment werd uitgevoerd om de hypothesen van dit model te testen. Hoewel de conclusies geen volledig uitsluitel gaven, wezen resultaten op een invloed van extreem negatieve recensies op waardering van de film. Er worden aanbevelingen gedaan om het model aan te scherpen voor toekomstig onderzoek.

Abstract

Entertainment product reviews are the subject of research in three fields; marketing research looks at review effects on movie performance and profitability, communication research inspects the rise of online consumer reviews, and a much smaller number of investigators focus on possible influences others' opinions can have on an individual's entertainment evaluation. The current study attempted to combine these fields by proposing an exploratory model that charts influence of critical reviews versus consumer reviews on post-viewing evaluation with involvement as a potential moderator, and tested this model for a short film in a controlled experiment. The results indicated polarised negative reviews can reduce enjoyment and involvement, though conclusions are tentative. Recommendations are made from the data to develop a confirmatory model that can unite the disparate fields of research.

1. Introduction

From an instrumental perspective, the consumption of experience products such as motion pictures is wholly unlike that of other goods. Their uses and worth are subjective and idiosyncratic for each member in their audience, and are hard to predict before consumption (Eliashberg et al. 2006). Also unlike most other goods, consumers cannot easily lay claim to a refund if an entertainment experience has failed to live up to expectations. For movies specifically, predictions of quality are made even harder by the way they are produced. Directors and producers, though often having preferences for certain actors and crewmembers, are usually not tied down to the same staff for different productions, and productions are often made by a company that was set up for the purpose of one movie (or a series of movies) and disbanded after completion (Hadida, 2009). Combined with the emotional nature and the need for each movie to feel like a different experience, this lack of persistent production staff stability means potential viewers cannot count on 'brand' names or even always names of project leaders to ensure they will receive the experience they pay for. Modern society offers a service that attempts to guide consumers towards more qualitative experiences; reviews written by critics (Holbrook, 1999). These reviews are often called expert reviews, and the emphasis the internet has started to place on user content generation has meant that these are now joined by the works of pundits voicing their opinions on goods, labelled as consumer reviews (Gao et al. 2006).

The current study was aimed at comparing the two styles of criticism in their effect on film viewers' opinions. To this end, a two-factor experiment was performed where individual subjects in four conditions were presented with reviews differing on two points; review valence and reviewer style, before watching the movie in question. Post-viewing movie evaluations formed the dependent variable for this study. The introduction to the current article will therefore focus on providing background to be able to posit hypotheses. Since critical reviews have been around longer than consumer reviews and have received more attention from researchers, evidence on effects exerted by the former, mostly garnered from studies with theatrical box-office revenue as a dependent variable, are discussed in the following section. Next, the few studies that tested for effects on enjoyment and evaluation are discussed. After looking at the rise of user-generated content trends, previous publications comparing the two styles of review - albeit not uniquely directed at movies - are inspected.

Following this background review is a section on theoretical models for review influences, and before the hypotheses and methods of the current study are outlined, potential moderators of review effects on post-viewing evaluation are reviewed.

1.1 Reviews and Box-office performance

The unpredictability of movie quality discussed in the opening paragraph affects their production as well. Just as potential viewers cannot be sure the movie they pay to see will live up to their expectations, investors and stakeholders in movie production do not know beforehand if their products will perform well in theatres, and are gambling capital on what they think are the best odds. In the past, these gambles have shown to sometimes lead to massive losses (Eliashberg & Shugan, 1997), while some movies with incredibly limited budgets went on to yield exponents of their investments (Kay, 2009). Obviously, the ability to predict the profitability of projects beforehand was considered lucrative, explaining the wealth of research on economic viability predictions for motion pictures. The extensive field review of Hadida (2009) showed that while 'film performance' - in this case signifying turnover for all mentioned dimensions except artistic recognition - was measured with numerous indicators, the dependent variable tested for the most was domestic theatrical box-office returns.

Though the discussion for the current article focused on influences of movie reviews, it is worth remembering that reviews are only one of the many variables exerting effects on movie performance and were found to be easily outweighed by different factors (Hadida, 2009). For example, one obvious factor that greatly influenced consumers' intentions of viewing a movie was its genre (Austin & Gordon, 1987). Reviews were disproportionately researched in light of their small effect sizes compared to other factors because they usually fall outside the range of movie producers' control, which incentivised research to facilitate the invention of strategies to maximise positive influences while minimizing negative influences (Eliashberg & Shugan, 1997). In previous studies, reviews tended to have shown greater effect sizes when tested for interaction effects. In the late 1980s, reviews had a more positive correlation with box-office results during the summer than other seasons (Litman & Kohl, 1989). Desai and Basuroy (2005) analyzed two datasets totalling 275 movies from the last decade of the twentieth century and found that while reviews had a significant positive correlation to the box-office results of movies that featured a movie star celebrity cast, they

could not rouse audiences to see (or choose not to see) movies that did not cast stars. A possible cause for this phenomenon forwarded by Desai and Basuroy (2005) was that potential viewers judge movies without celebrity casts as inferior - again showing that reviews are not primary motivators - and movies that do feature celebrities are still seen to differ enough in quality to warrant review effects.

One prominent question recently raised by researchers called the nature of the relationship between critical reviews and box office revenue into question; do critics influence their audience, causing changes in viewer numbers, or do they simply reflect the crowd's interest in seeing a movie, acting as predictors of revenue? Eliashberg and Shugan (1997) had started to disentangle the problem by comparing critical rating consensus for 56 movies to box office revenue across the first eight weeks of their theatrical run. Although their results suggested professional movie reviewers act as predictors of consumer interest, results were only obtained at aggregate level by examining movie statistics. Further, they induced their theory of the critic as indicator from the finding that box-office revenues correlated significantly with the percentages of positive and negative reviews only in the time period after the movies were in theatres for more than four weeks. Though this phenomenon was corroborated by Plucker et al. (2008), it does not present strong theoretical support for their conclusion, relying on the effect of word of mouth to smother or heighten moviegoers' interest to the level predicted from the critics' evaluations. Finally, the expansion by Basuroy et al. (2003) of Eliashberg and Shugan's (1997) study, having increased the sample from 56 to 200 movies, brought results that suggested negative reviews negatively influenced box-office results sooner after release than the previous study, adding further doubts to the notion that critics only show predictor correlations with box office results.

Gemser et al. (2007) reported differences between Dutch art-house and mainstream movies in critics' relationships to box office performance. Mainstream box office results were found to be predicted by critical reviews, while an influencer effect was apparent on art-house movie performance. Coming to a similar conclusion as Desai and Basuroy (2005), they judged that art-house movies showed greater influencability from critics because their lack of mainstream appeal and budget meant critics' reviews simply spread awareness. In contrast to the latter study, however, Gemser et al. (2007) found that the valence of the review did not matter significantly.

Reinstein & Snyder's (2005) attempt to separate predictor correlations from influences approached the issue differently. Instead of aggregating data from a large number of critics, the sample was limited to two prominent reviewers, Gene Siskel and Roger Ebert, whose reviews were televised nationally in the United States. Operating from the viewpoint that if any critics had the ability to influence box-office results, it would be Siskel and Ebert, opening weekend revenues from movies whose reviews were broadcast during their first weekend - the show usually aired on Saturday morning - were compared to those of movies whose reviews were shown after this pivotal interval. Since reviews disseminated after the opening weekend could not influence returns during it, the critic as predictor was separated from the critic as intention influencer. While the pure influencer results were only barely significant for the entire sample of movies, dividing them into genres and release width (the number of screens showing it across the country) yielded significantly positive results for dramas and narrowly released films (Reinstein & Snyder, 2005). Evidence corroborating the pure influencer effect is hard to obtain since few researchers have taken explicit measures to separate it from predictor correlations. The current body of research suggests that critics likely both predict their audience's interest as well as influence it.

1.2 Impact on evaluations

The inconclusive results caused Eliashberg and Shugan (1997) to call for individual level experiments. Presumably, this would also mean moving away from box office revenue as a dependent variable. Unfortunately, though research on the topic of movie reviews is rife with marketing studies explicating the role reviewers play in increasing or reducing box-office revenue, significantly less attention has been paid to the possible influence critics might play in the generation of consumer enjoyment during a screening and their opinions on a movie afterwards. Correlational evidence linking critics' review verdicts to viewer ratings was provided by Boor (1992). As well as showing significant correlations between the viewers' ratings and the six investigated critics' verdicts, the critics displayed high agreement amongst themselves. It is intuitively likely that this latter agreement does not stem from any influence, but simply shows like-mindedness. However, the same cannot easily be said for the relationship between viewers and critics. Burzynski and Bayer (1977) conducted a short experiment where they made waiting movie patrons overhear negative or positive comments from confederates seemingly coming from a previous screening. Consequently, patrons who overheard positive comments awarded the movie with a more positive evaluation - measured

as a grade 1 through 10, where "1= *Terrible*, 10= *Excellent*" (page 217) - than patrons who had overheard negative comments. Differences between either group and a group that was not presented with prior information did not reach significance. Burzynski and Bayer (1977) reported that the movie patrons 'conformed' their evaluations to them. The fact that moviegoers' evaluations can be altered by strangers simply voicing their unfettered opinions within earshot lends credibility to influences exerted by an expert review.

Wyatt and Badger (1984) performed a controlled laboratory experiment to find such an influence. A film was chosen that garnered an equal number of positive, mixed and negative reviews when it was released. A basic review was then written that allowed the valence of the critic's opinion to be altered by changing the adjectives in each sentence. This produced three reviews that were identical in structure and neutral information about the movie, but differed in the critic's positive, mixed or negative judgements on elements such as the writing, acting and cinematography. Subjects were shown the review as part of a large questionnaire to obfuscate the aim of the study. Results indicate that negative and positive reviews caused significant differences in subjects' evaluations, though only the negative treatment made responses differ from a control group while only the positive treatment made responses differ from respondents who read a mixed review (Wyatt & Badger, 1984). The respondents in the control condition rated the movie highly, which explains why the positive treatment did not differ significantly from the control condition. This illustrated the importance of the entertainment content used in such an experiment. The dearth of research on this topic means there were no proven guidelines to facilitate material selection for the current study.

The relationships between expert judgments, consumer evaluations and popular appeal of movies were modelled by Holbrook (2005). Using the Internet Movie Database (www.imdb.com) to test a sample of 219 movies, the number of ratings were taken as an estimator of popular appeal while the average total rating for each movie was noted as its consumer evaluation. Results supported the notion that while consumers' ratings correlated positively with expert judgments, a movie's popularity (volume of ratings) was only weakly related to both expert judgments and consumer evaluations. The explanation forwarded for this model was that although consumers knew what makes a movie critically 'good', they could still enjoy movies that did not live up to this standard for other reasons. Holbrook

(2005) could not provide evidence for this explanation because, as in Eliashberg & Shugan (1997), results were aggregates and did not include personal queries.

A promising avenue of research shifted the focus of review influences to the reviewer. In the area of professional criticism, influences reviews can have are moderated by the identity of the critic. d'Astous & Touil (1999) performed a complex multifactorial experiment with a total of six factors spread between and within subjects. Subjects read fictional summaries of critics' opinions and were informed of his/her verdict history on other films and whether a consensus was reached by different critics on the current movie. While the subjects did not actually see any movies themselves and were not presented with full reviews, this experiment showed that audiences engage in attributional thinking of a critic when judging an unseen film by this critic's verdict. The generalizability of these results is questionable considering it is unlikely that readers research each critic's history when simply reading a review for a movie they are interested in. It is left to the popularity of critics to promote attributional thinking, though it is likely only a very select few are recognized by name by the general public (e.g. Siskel and Ebert).

1.3 Reviews in the information age

In recent years, internet culture has increasingly been geared to the creation of content by each individual user, leading to the rise of social media (Kaplan & Haenlein, 2010). The new emphasis on user-generated content, vastly different to traditional media that only allowed a select few to voice their opinions, has upset the previous 'cultural hierarchy' (Holbrook, 1999). This had major consequences for the creation as well as the consumption of reviews (Gao et al. 2006). The greater number of reviewers combined with the ease with which data can be aggregated (like the data used in Holbrook, 2005) led to a greater focus on numerical and ordinal ratings such as the ubiquitous one to five 'stars' (Tsang & Prendergast, 2009). The cost for the usefulness of reviews is that nuances accompanying the rating in the review's text are lost during aggregation. Tsang & Prendergast (2009) noted that though consumers at least partly base their intention to see a movie on a review's rating for a movie, this effect can be overridden by its text. The text influenced perceived trustworthiness, especially for positive ratings. This meant that there still was a need for individual reviews explaining why a movie is (not) recommended. A similar effect was not found in freeware software research, where the review valence of either critical or consumer reviews was not

nearly as influential as the volume of posted reviews in predicting downloads (Amblee & Bui, 2007).

Researchers agree for the most part that differences exist between consumer and expert reviews. Gao et al. (2006) asserted that "Prior literature generally assumes that consumers report their private observation in product reviews" (page 2), and found evidence for the notion that the use of private information (for assertions in consumer reviews) was reduced when more sources of public opinion were introduced, holding especially for expert opinion. Dellarocas and Narayan (2006) found evidence for a mild polarization effect in consumer review ratings, creating a bimodal (U-shaped) review spectrum where consumers lean either towards positivity or negativity. Conversely, critical reviews are usually unimodally distributed and significantly intercorrelated (Boor, 1992). Book reviews on Amazon.com allowed for insight into differences in behaviour of reviewers with different levels of experience (Shen, 2009). Inexperienced reviewers conformed to the U-shaped distribution, attempting to stand out from the crowd by posting extreme reviews that often differed from consensus. A selection bias was also found whereby experienced reviewers chose to write about less popular books while inexperienced reviewers opted for more high-profile books.

Plucker et al. (2009), on the basis of a comparison between experts and several different samples of movie novices, suggested discriminating between "novice critics" (page 476), amateur reviewers posting on movie forums (e.g. IMDB.com and boxofficemojo.com) and complete laypeople. Laypeople, by this definition, do watch an average amount of movies but usually refrain from writing reviews. This crowd does weigh in on average ratings for movies by giving grades to movies they have seen, especially if it exceeded or failed to meet expectations by a considerable margin (lent credence by the bimodal rating distribution discussed previously). Differing rating styles do not preclude significant correlations between professional critics and lay ratings; Plucker et al. (2008) report an r of .71 between ratings from critic score compilation site Metacritic.com and ratings on IMDB.com and Holbrook's (2005) comparison of published critics and the same crowd on IMDB.com yielded an r of .85. Despite their high intercorrelation, most review compilation sites still differentiate between expert and consumer reviews. Their combined presence interact to form new factors in consumers' content selection. Wang (2005) showed subjects movie information web pages that separately displayed consumer and expert ratings, and tested for credibility of either

review type and attitude towards the movie. While expert ratings were judged to be more credible and both types had the ability to raise pre-viewing attitude, subjects had stronger intentions to see a movie when they placed more credibility with consumer ratings and a consensus was found between the types.

1.4 Toward a model of review influence on evaluation

The changing landscape of movie reviews warranted a revisit of investigations of direct influence of reviews on movie evaluations. Because most of the literature discussed previously assumed a marketing or economics perspective, very few psychological theories were used to model effects. The current section focuses on elucidating the concept of a viewer's evaluation of a movie by examining enjoyment. Likely theoretical mechanisms for reviews to influence enjoyment are reviewed, as well as possible moderators.

The inclusive definition of enjoyment maintained in the current article is outlined in Vorderer et al. (2004). It is described as a reaction process, contingent on cognitive, affective and even physiological processes, that is the largest driving force in entertainment selection. Enjoyment does not rely solely on hedonistic gratification, but also on eudaimonic rewards. Lastly, enjoyment is not just a product of the entertainment content presented, but one born out of the interaction between the viewer's motives, mood state, stable personality characteristics and the entertainment experience. Presumably, people employ a naive version of this concept when providing an evaluation for a movie. Movie evaluations are usually summarized in ratings (e.g. Holbrook 2005) or measured with a list of polar adjectives (Wyatt & Badger, 1984), relegating the concept to an overall judgment of quality. None of the publications mentioned either briefing their participants on how to judge what they have seen and which aspects to consider, or gauging from what perspective their judgements came. Participants could be displaying several evaluative standpoints; Holbrook's (2005) popular appeal, for example, differs from consumers' "ordinary evaluations", though both measures originated from the same population. What is creatively excellent might not be entertaining to watch. The popular appeal likely reflects a hedonistic factor in movie evaluations, where the content is popular simply when it is fun to watch (Vorderer et al. 2004). Ordinary evaluations, the consumer's conscious verdict, probably consist of hedonistic and critical elements, where factors such as acting talent, cinematography and narrative all become more important. Here, the novice reviewer's lack of knowledge on how to judge quality comes into play, evident in

the use of 'private' information and that of professional critics for a consumer review discussed earlier (Gao et al. 2006). This speculation places (consumer) enjoyment of a movie at least partly outside of the elements of review by a professional critic, since enjoyment can be - and sometimes is - dissociated from critical consensus (Holbrook, 1999). When the reviewer is seen as an influencer of readers' opinions, the readers could be said to compare their own (novice) verdict based on private information with the critic's analysis and append weights similar to those discussed for social judgment in Fiedler et al. (1996-I) to both sources to arrive at a final verdict. When the critical judgment is deemed more advanced and well-grounded than the readers', greater weight will be placed on the review evaluation.

The difference between a critic's evaluation and that of an ordinary consumer can likely be brought back to their mindset while viewing, often called their mode of reception (Wirth, 2006, Vorderer, 1993) or framings (Liebes & Katz, 1986). In the latter, the objective of the study was to compare the ways groups from different cultures talk about a fictional television series' occurrences, though the framings can be extrapolated to the current situation. Framings were either critical, treating the series as an artificial construct or story, or referential, where the series was treated as having bearings on real life and events in it had real-life analogues. Referential framings place the events in the viewer's world, allowing full emotional engagement. Critical framings, on the other hand, create distance between the portrayed events and real life, allowing viewers to follow the content without getting emotionally involved. In Vorderer (1993), the critical viewing mindset is defined as a mode of reception that does not allow for much involvement. Viewers who are fully involved with the content are engaged in the events and allow themselves to be stirred emotionally. These modes are not fully dissociated, and viewers can also switch between them. An exogenous example of such a switch is when actors fail to convince a viewer, and the viewer starts to notice that she is watching an artificial event. It follows that professional critics are likely to 'see the strings' of a production more often and sooner than regular viewers due to their experience. This would lead them to be less affectively involved (Wirth, 2006) and experience the emotions more naive viewers go through on a different level or not at all. So, while one of the most important products of showing a motion picture is viewer enjoyment, it is usually not the only element mentioned in a critical review. Involvement can also occur on a cognitive level, of course, and in the field of consumer research (Holbrook and Hirschman 1982), similar modes of reception are referred to as information processing (critical) and experiential

(affectively involved) to remove confusion. The current study uses the term critical to refer to the more cognitively involved mode of reception, and experiential to denote the more affective mode.

Despite the complexity of reading and comprehending a review, participants need not be consciously aware of any hypothesized effect for it to take place. Discoveries in the field of evaluation priming research have consistently shown influences of primes on subsequent judgments while participants explicitly reject the prime's information (Fiedler et al. 1996a). Individuals also displayed changes in behaviour because of certain information even after that information (or its source) had been discredited through an experimental debriefing paradigm (The perseverance phenomenon: Ross et al. 1975, see also Gilbert et al. 1990). The influence of source credibility was observed for the current situation. Another example can be found in presupposition effects; effects exerted by the phrasing of questions that contaminate evaluative attitudes without the individual's awareness (e.g. Fiedler et al. 1996b and Wegner et al. 1981). Many of these effects are related to the way people handle propositions, often referred to as the Spinozan model (Gilbert et al. 1990). This model states that, when an individual is presented with information, it is automatically taken as true through the act of comprehension (Koehler, 1991). Only after the statement has been comprehended can it be 'tagged' as false, a process that can be interrupted.

Though these phenomena were usually researched with fundamental-level experiments with short intervals between treatment and measurement, the results were applied to news message influences (Wegner et al. 1981), and can likely be extrapolated to the current situation. Thus, while subjects might deny being affected by the valence or style of a review, a critical or referential phrasing of an evaluation could be seen as a presupposition for judging entertainment content. If this is so, effects should still be measurable whether or not the subject is consciously aware of the manipulation. Similarly, a review's judgment can logically only be discounted while reading the review - for example because the reader is not impressed by the strength of its analysis - or afterwards, when the reader experiences the movie differently than the reviewer described. This means that the review is taken as true at least as long as it is being read (Koehler, 1991) and likely beyond that to the viewing itself and possibly even when the reader is asked to give her own verdict. A possible explanation for effects found in this field - and a possible confound for the current study - is that the reviews cause a demand effect in readers (Fiedler et al. 1996a) whereby they take the valence as an

indicator of how to judge the content as the experimenters wish it (McCloskey & Zaragoza, 1985). In this case, however, conscious awareness of the manipulation is likely prerequisite.

Because most entertainment media content involves human or anthropomorphised characters, it is not surprising that enjoyment is largely dependent on the viewer's connections with them (Vorderer et al. 2004). Many theories have been forwarded to chart these connections, such as parasocial interactions and relationships (Horton & Wohl, 1956), identification (Cohen, 2006), and affective disposition theory (Zillmann, 1994). Affective disposition theory (ADT) places emphasis on the formation of dispositions toward characters and their influence on the experience of empathy for the protagonist. In this view, trait empathy, the relatively stable individual factor that governs the propensity to feel for characters, is very important to engage with dramatic content on an emotional level (de Wied et al. 1994). Greater empathizers noted higher enjoyment of dramatic motion pictures, although their experienced emotional valence during stressing times in the narrative was more negative. These peaks and troughs suggested that individuals with higher trait empathy were more engaged in the story, indicating an experiential mode of reception.

The previous literature in marketing research suggests at least a small influential effect of the appearance of critical reviews on post-viewing evaluations. Though few publications attempted to define the relationship between consumer reviews and evaluations, the extant body of work indicates the number of consumers reviewing a production is more important than the valence of single reviews. Neither research subject presents a strong theoretical model for the mechanism of influence of reviews on evaluations in individual cases. The current study forwards such a theory and tests several hypotheses to provide evidence to support it.

Critical reviews present a reader with a confident critical analysis of the quality of a motion picture. Conversely, consumer reviews present a (relatively) novice viewer's opinion, fuelled by the writer's urge to speak out on a movie. Critical reviews are hypothesized to confer unto readers a critical or analytical reception mode, causing viewers to, for example, notice acting talent and reflect on the narrative beyond an experiential mode of reception. While this mode might change during the viewing, an enquiry to the quality of the experience in a post-viewing questionnaire will likely cause viewers who read a critical review to think about the movie in a more analytical and emotionally distant frame. If readers feel the

analysis was suitably intelligent and complex they might conform their own opinions to the review valence. Consumer reviews might also influence the post-viewing evaluations, but will not decrease involvement over viewers that have not read a review. However, if the reader does not trust the review, no difference to a control group will be found. Finally, trait empathy can moderate the effects of review valence and style, with high empathizers evaluating a movie more positively. A conceptual model of this tentative theory is displayed in figure 1.

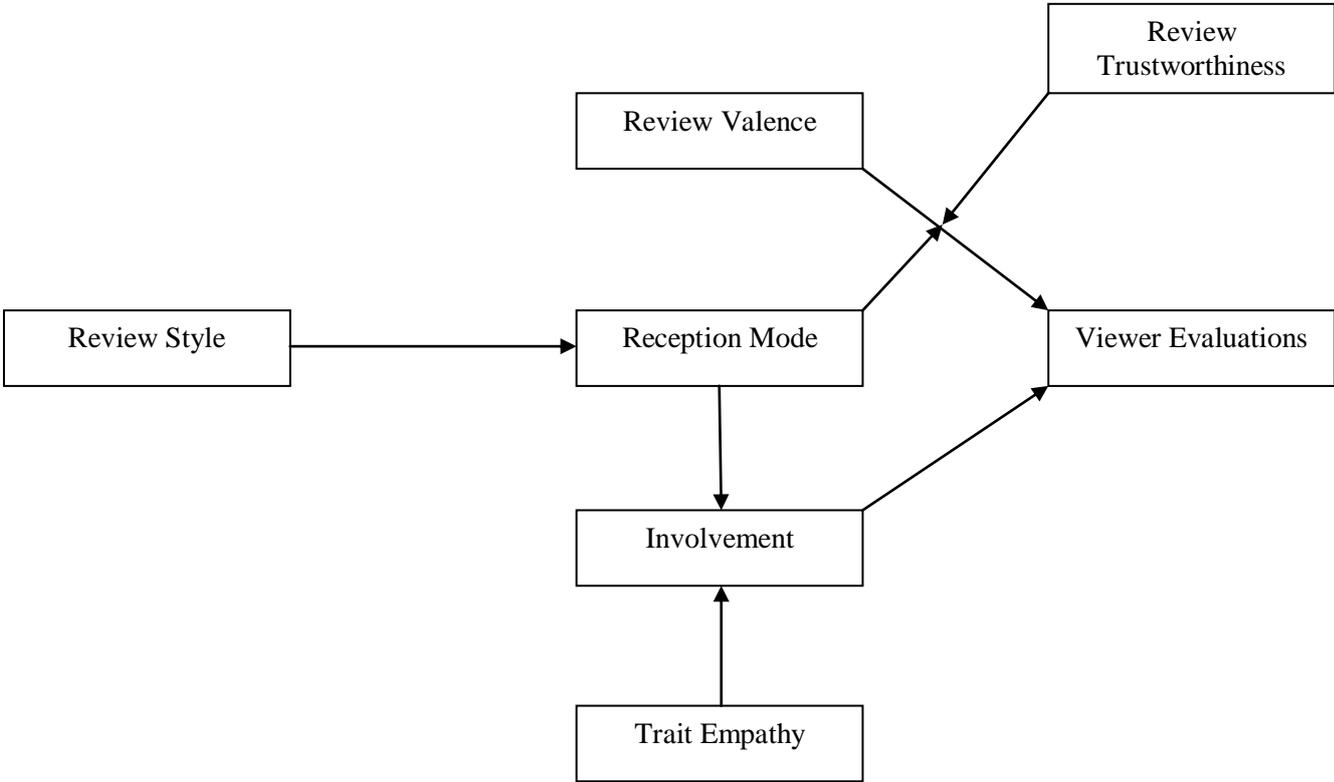


Fig. 1; Conceptual model of review influence through reception mode and evaluative framing.

The following hypotheses can be garnered from these predictions:

- H1: Critical reviews make readers conform their evaluations to the review valence.
- H2: Critical reviews decrease measures of affective involvement while increasing agreement on analytical elements discussed in the review compared to consumer reviews.

H3: Consumer reviews make readers conform their evaluations to the review valence.

H4: An interaction effect of style and review valence shows greater conformity to the critical review valence than to the consumer review valence.

H5: Trait empathy is positively related to involvement and viewer evaluations.

2. Methods

2.1 Participants

The sample consisted of 90 students of a Dutch university, majoring in either psychology (63%) or communication studies (36%). Participants were between the ages of 17 and 25, while the mean age was 20 years. The majority of the sample was female (76%), with only 21 males participating, though this is reflective of the student population in this field. The distribution of participants across conditions - detailed in subsequent sections - and gender is displayed in table 1. 69% of the sample was in their first year of college. The sample consisted of Dutch and German students, though their relative proportions were unknown. The possibility exists that the German students found the Dutch treatment stimuli more difficult to comprehend. However, any confounds for the study that could result from the different nationalities were ameliorated by the randomization of participants. Furthermore, students are expected to be proficient in the Dutch language. All respondents were given course credit for their participation, and an extra incentive was added by allowing full participants to enter in a raffle for a gift certificate worth € 50. When tested for differences on demographic variables with ANOVAs, the conditions show near-significant divergence. Differences in gender ($F(1, 88)=3.43, p=.067$) and age ($F(1, 88)=3.13, p=.081$) remain just outside significance for the two review styles and the control condition.

Table 1: Distribution of participants across conditions

	Critical Positive	Critical Negative	Consumer Positive	Consumer Negative	Control	Total
Male	8	5	4	2	2	21
Female	12	15	15	17	10	69
Total	20	20	19	19	12	90

2.2 Stimuli (Motion Picture and Treatment)

The current study performed measurements using an individual-level experiment. One film was chosen as a target experience. Following an investigation of the literature, several clues to guide the selection of an appropriate movie became apparent. Though many studies (e.g. de Wied et al. 1994, Wyatt & Badger, 1984) edited their material to shorten it, an originally short film - roughly 30 minutes in length - was found for the current study. This had several advantages: it reduced the testing time of the experiment (1) while still presenting a full narrative that had not been artificially manipulated and appeared as the director had intended it (2). Finally, it had not received the marketing push or (wide) cinematic release typically seen for feature films in spite of 'Hollywood' production values, meaning few participants had heard of it even though it was released in 2010 (3). The production was titled *'I'm Here'*:

"I'm Here is a 30-minute love story about the relationship between two robots living in L.A. The film is written and directed by Spike Jonze. Andrew Garfield and Sienna Guillory are in the lead roles, and the soundtrack includes original music by Sam Spiegel and original songs by L.A.-based art musician Aska Matsumiya and other emerging musicians."

- About section of 'www.imheremovie.com'.

I'm Here was chosen for the current study for a host of reasons; it was written and directed by an acclaimed and experienced director; it conferred roughly the same three-act structure apparent in most full-length feature films; the main characters were played by actors whose faces were never revealed to prevent enjoyment contamination through visual recognition (for example through the mere exposure effect described by Zajonc, 1968); it was not well-known to the general public; its production value was higher than most short films; it portrayed a love story with added elements of sacrifice and racial discrimination (and so could be considered a dramatic movie, maximizing review influence according to Reinstein & Snyder, 2005); and it was described by third parties as having a quirky atmosphere with art-house elements (recommended by Gemser et al. 2007). A disadvantage of using *I'm Here* was the consensus of positive reviews it received. Almost three thousand IMDB users rate it an average 7.9 out of 10, while review compilation site rottentomatoes.com scores it 4.1 out of 5 with a 90% approval rating. The positive evaluations it has received could present confounds into the data. The current study's sample rated the movie an average of 6.7 out of 10. None of

the respondents reported having seen the film before they participated, although one possible participant was rejected because she had performed an in-depth analysis on *I'm Here* for a educational project.

In order to test the hypotheses, reviews were needed from both a critic's and a consumer's perspective. Furthermore, for both of these styles, positively and negatively valenced reviews were required. To preserve internal validity, a fictional review was written for each of the four permutations. A similar strategy to Wyatt & Badger (1984) was pursued in generating reviews. This strategy entailed writing two reviews ('critical' and 'consumer') that could be manipulated to be either positively or negatively valenced, while remaining as similar as possible in objective information presentation, review length and cinematic elements discussed. Considering the possible confounds (Tsang & Prendergast, 2009), no numerical ratings were included with any review.

The critical reviews were split into four paragraphs and featured the reviewer's (fictional) name, a screenshot showing the main character and a list of details such as staff credits, distributors and date of appearance. The reviews first mentioned the writer/director who was described as either 'critically acclaimed' (positive valence) or 'relatively unknown' (negative valence). Both reviews mentioned the theme of the film, summed up the plot, scrutinized the visual effects, casting choices, noted its short length and relative obscurity, and closed by either calling the production 'a flashy mess' (negative valence) or 'an undiscovered pearl' (positive valence). Each sentence's adjectives and quality signifiers were polarised to underline the final verdict, though care was taken to retain a realistically professional tone. The positive review was 164 words in length while the negative review numbered 165. Both reviews were printed in a lay-out that very closely mimicked a Dutch professional critic's magazine (www.filmkrant.nl), copying its logo and front-end. Though this publication never actually reviewed *I'm Here*, the reviews were given a fictional volume number and a date of publication.

The consumer reviews were presented as user comments on a film aficionado forum (www.moviemeter.nl). The reviews were embedded in mock-ups of a movie's information page, though average rating information was not included. The top half of the page contained the logo of the forum, the movie's title and information on genre, length and staff, as well as a screenshot of the main character and the number of movie theatres screening it (none). To

confer more realism, Kijkwijzer viewer advisory ratings (similar to MPAA ratings) were added, although *I'm Here* was never rated. The rating symbols used signified a minimum age of 9 years and noted elements of racial discrimination. The review was placed in a text box resembling a forum post, contained a standard avatar picture and the moniker 'movielover14' and was time-stamped.

To retain internal validity, the objective information (director and actors names, movie length, basic plot) was held constant with the critical reviews. The two review styles were differentiated by a number of elements. The consumer review contained many self-references from its author (e.g. 'I've just seen *I'm Here* and I can highly recommend it' in the positive consumer review). The consumer review was also not formatted into paragraphs, resulting in a block of text. Use of quality signifiers were more experiential in the consumer review; the negative consumer review described the special effects employed as being 'ugly', while the negative critical review remarked: "Visually, the movie fails to impress". The writing was simpler for the consumer review, meaning the vocabulary was more limited and words contained less syllables. To support this assertion, the review texts were entered into the Douma readability formula (Burger & de Jong, 1996), a Dutch version of the Flesch reading ease formula (Flesch, 1951). The Douma returned respective scores of 50 and 48 for the positive and negative critical reviews and 75 and 72 for the positive and negative consumer reviews. Since lower scores from this formula entail decreased readability and greater text complexity, the critical reviews were determined to be more complex. Comparing the scores using a legend, the consumer reviews scored 'fairly easy', corresponding to the reading level required when children start high school. The critical reviews score as 'difficult', corresponding to reading level required in later years of Dutch high school's upper echelon. On the basis of this pre-test the review styles were concluded to differ appreciably in their complexity. A side effect of the simplification for the consumer reviews was that they had to be larger to be able to include the elements shared with the critical reviews. The positive consumer review contained 229 words (65 words longer than positive critical), while the negative review consisted of 238 words (73 more words than negative critical).

The four reviews were pretested with 6 individuals that did not participate in the main study and were not informed of the purpose of the study. The participants were asked to provide a numerical rating between one and ten that they judged the reviewer to have given the film. The positive critical review was judged as presenting the movie with a grade of

7.8/10, while the negative critical review received an average of 3/10. The consumer reviews were deemed to be graded 8/10 for the positive and 2/10 for the negative review. The conclusion was drawn that readers are able to discern the negative from the positive reviews.

The control group was shown a purely informative page, based on the lay-out of IMDB.com. This page featured, as the review pages in the other conditions, basic information on the movie's length, genre, year of production, director and cast list. The storyline is briefly summed up with a picture of the protagonist, but care was taken to remove ratings or any positive or negative comments.

2.3 Procedure

The experiment involved a quasi-experimental design with two between-subjects factors set by conditions (review style and valence) and one between-subjects factor (trait empathy) that was uncontrolled. Participants were randomly placed in one of five conditions and tested individually; two review styles and two review valences for each, and a control group. The testing location was an office room. After welcoming participants, they were minimally briefed. They were told they were going to see a movie that was 30 minutes long and that they would fill in a short questionnaire afterwards. Next, the participants filled in a written consent form. An attempt was made to obfuscate the goal of the study by presenting the reviews as merely 'information'. Participants were informed the movie was not well known and that the pages were the only source of information that could be found. The top of each page was titled as follows: "The following information was taken from..." followed by the name of the website the lay-out was drawn from.

After participants indicated having read the page the reviews were removed and the participants were placed in front of a 26" CRT flat-panel television. To minimize distractions, they were given a pair of headphones and the television remote to control the volume. English hearing-impaired subtitles were used. The room was darkened and the experimenter sat on the opposite side of the room while participants watched the movie. After the viewing, the participants were presented with a six-page questionnaire. Upon completing the questionnaire, they could leave contact information to compete in a raffle for a gift certificate. Before thanking them for their participation, subjects were asked if they had any questions. To minimize contamination in the student population, participants were not explicitly told the purpose of the study if they did not ask about it. Each session lasted around 45 minutes.

2.4 Questionnaire

The study's questionnaire was six pages long and split into 6 sections. The first section asked participants to grade the movie they had just seen on a scale of one to ten. No information on how to make this judgement was included. This section was placed before the others to avoid 'carry over' and 'backfire' effects (Tourangeau & Rasinski, 1988) from other items in the survey. The second section contained 15 adjectives, and participants were asked to signify the fit each adjective had with the film on a 7-point agree/disagree scale. The adjectives - translated to English - are; Interesting, fun, entertaining, long-winded, subtle, beautiful, poetic, pretentious, artful, short, sad, pathetic, funny, serious and exciting. Six of the adjectives were used in the reviews to describe the film (i.e. interesting, subtle, beautiful, pretentious, artful and short). The third section of the questionnaire asked participants whether they agree with 9 propositions concerning the way they had seen the movie. These propositions included both critical - 'this movie had good acting talent' - and experiential - 'I felt connected to the fate of the characters in this film' - statements, and agreement was gauged on a 7-point agree/disagree scale. The fourth section was a trait-empathy scale, based on the multidimensional approach and scales devised in Davis (1980) and refined in Davis (1983). Three out of the four scales developed were translated to Dutch; Perspective Taking, Fantasy and Empathic Concern. The fourth scale from Davis (1980), Personal Distress, was excluded. This scale was aimed at gauging reactions to emotional emergencies and was deemed unsuited for the usually calm activity of watching a film. For each of the remaining scales, seven items were placed in the questionnaire, yielding 21 statements with 7-point agree/disagree answer options. The fifth section of the questionnaire consisted of seven statements on the review read. Participants were asked whether the review was easy to understand, easy to read, fun to read and trustworthy, whether they thought the reviewer thought his review through, and finally asked if participants often read reviews like the one they read and whether they usually agree with them. The final section had demographic items but also asked participants if they had seen the movie, what rating the reviewer would likely have given it, and what the reviewer's name was. The final item in the questionnaire was a manipulation check, allowing participants to guess what the purpose of the study was. This item was graded as incorrect, partially correct, or correct.

2.5 Data-analysis

The items in the questionnaire were averaged to form construct variables. Three construct variables were created. The adjectives from the second section were used as an aggregate evaluation scale comparable to the numerical evaluation in section 1. Statements about the way the movie interacted with the viewer from section 2 formed an involvement variable, while the empathy-related items adapted from Davis (1980) were combined into a single trait-empathy variable. First, the items were content-inspected to determine whether they needed to be recoded. Item list were then tested for reliability using Cronbach's α , and items that lowered the α were recoded or removed. Next, principal components analysis determined the number of constructs involved in each scale. The recoded and refined scales were averaged to create the tested variables. The adjective evaluation scale consisted of one major principal component and showed a satisfactory α of .79, although the items 'sad' and 'pathetic' were removed, leaving thirteen items. The nine statements about the film's influence showed very good interreliability ($\alpha = .88$), and were all influenced by one major component. Despite the intentions of using a part of this scale for an experiential variable and the other part for a critical variable, all items measured a construct that had the highest loadings on experientially-oriented items. No α improvements were possible if the critical items were removed, so all items were averaged into the involvement scale variable. Finally, the items measuring trait empathy pointed to a small number of components. While two smaller components were ostensibly related to the different subsections of the scale, one large component correlated significantly with all items. One item (translated as 'I often fantasize about things that could happen to me') detracted from interreliability and was removed from the scale. The final scale showed an α of .82, and all items were averaged into a single trait-empathy variable. The review-related items from section 5 were used separately.

The dataset was inspected for outliers using boxplots. Two participants gave extreme responses to the questionnaire's statements on multiple variables. Both participants were in the critically negative review condition, and one of these failed to complete the full questionnaire. The cases were removed from the dataset. Lastly, the main variables were tested for skewness and kurtosis in their variance. Adjective evaluation, trait empathy and involvement all showed skewness and kurtosis values between -1 and 1. Numerical evaluation had the highest values for both tests (skewness: -.73, kurtosis 1.52), though they were not

extreme. The variables were close enough to being normally distributed to allow for parametric tests.

Data-analysis was performed using PASW Statistics 18. Tests on single factors were performed as t-tests or ANOVAs where applicable. Multiple concurrent model tests were performed as univariate analyses of variance with linear regressions. R^2 of the dependent variable was taken as an indicator of the independent variables' explanatory power. Because the model forwarded in the introduction was untested, exploratory structural equation modeling (SEM) was performed to make inferences on the model's value for the current situation. SEM tests were performed using Amos 18.0. The choice for model fit indices was guided by Holbert and Stephenson (2002), and included the following: Chi-squared with degrees of freedom and their ratio (X^2/df) as a relative fit index, the Tucker-Lew Index (TLI) as an incremental fit index, and the root mean squared error of approximation (RMSEA) and the standardized root mean squared (SRMR) were used as absolute fit gauges. Models with good fit normally have a TLI above .95, a SRMR below .09 and an RMSEA below .06 (Hu & Bentler, 1999). Standardized regression weights (β) were scrutinized for individual effects. Though the sample size was not optimal for SE modeling, the model's simplicity and lack of latent variables made SEM a viable option for analysis.

3. Results

Table 2 displays the distribution of the average numerical evaluation, adjective evaluation, involvement score and attributed reviewer grade across conditions with standard deviations in parentheses. The average numerical evaluation of 6.7 out of 10 for the sample is reflected in the average the participants in the control condition report (6.8). The grade the reviewers were attributed to give the movie (perceived reviewer grade) shows a small divergence from the pre-test; the critical reviews were perceived as more moderate (7.6 for the positive review and 4.3 for the negative review) than the consumer reviews (8.9 for the positive review and 2.6 for the negative review). Unsurprisingly, a significant main effect of the review valence ($F(1, 71)=276.04, p<.001$) was found when estimated reviewer grade was used as dependent variable. The main effect of review style did not reach significance ($F(1, 71)=.47$), though the interaction effect of the two manipulations did ($F(1, 71)=25.66, p<.001$), lending credence to the notion that critical reviews were more moderate. The consumer positive review was evaluated as being more positive than the critical positive review

($t(37)=3.49$, two-tailed $p=.001$) and the consumer negative review was more negative than the critical negative review ($t(34)=3.65$, two-tailed $p=.001$).

Table 2; Means (and standard deviations) of variables of interest across conditions.

Review Style	Review Valence	Numerical Evaluation ¹	Adjective Evaluation ²	Involvement ²	Perceived Reviewer Grade ¹
Critical	Positive	7.0 (1.49)	4.6 (.90)	4.6 (1.13)	7.6 (1.40)
	Negative	6.6 (.87)	4.3 (.69)	4.4 (.90)	4.3 (1.53)
Consumer	Positive	7.1 (1.12)	4.5 (.63)	4.7 (1.04)	8.9 (.76)
	Negative	5.9 (1.55)	4.0 (.72)	3.9 (1.24)	2.6 (1.20)
Control		6.8 (1.48)	4.6 (.66)	4.6 (1.05)	n.a.

¹: Grade between 1 and 10.

²: 7-point scale.

Before the hypotheses are tested, the uncontrolled factors are reviewed. Trait empathy, fortunately, did not differ significantly between the conditions ($F(1, 87)=.03$ for review style and $F(1, 87)=.06$ for review valence). The manipulation check, where participants attempted to sum up the research objective, bordered on significance at .05 level for the review style used ($t(74)=-1.84$, two-tailed $p=.071$), with differences in mean suggesting the consumer review readers might have been more aware of the purpose of the study. Overall, 56% of the respondents in the treatment conditions were aware of the manipulation when they responded to the item (the last item in the questionnaire), with another 7% suspecting something similar. While a number of these are likely individuals who genuinely saw through the obfuscation, an indeterminable portion could infer the goals by looking at the previous sections of the questionnaire, and so were not aware of the purpose of the study when filling out most of the questionnaire.

Significant differences were found when inspecting the results of the item measuring trustworthiness of the reviews; a two-factor analysis of variance yielded main effects of both review style ($F(1, 75)=5.29$, $p=.024$) and review valence ($F(1, 75)=5.08$, $p=.027$), though no

interaction effect was found. The critical reviews were judged to be more trustworthy than the consumer reviews, while the positive reviews were perceived as more credible than the negative reviews.

A similar pattern was found on the statement 'the review was easy to understand' (main effect for style: $F(1, 75)=9.57, p=.003$, main effect for valence: $F(1, 75)=6.43, p=.013$, no interaction effect), with the consumer and positive reviews rated higher than the critical and negative reviews respectively, though the statement 'the review was easy to read' did not confer a main effect for the review style (main effect for valence: $F(1, 75)=4.89, p=.030$). Readability was higher for the positive than for the negative reviews. When the participants were asked whether the reviews were fun to read, the style did not matter significantly. The positive reviews were considered more fun to read than the negative reviews ($F(1, 75)=36.61, p<.001$). The same result was discovered for the statement that 'the reviewer thought about his opinion' ($F(1, 75)=22.24, p<.001$). When narrowing the confidence interval to an α of .10, respondents in the critical treatments stated they read reviews like this often significantly more than respondents in the consumer treatment conditions ($F(1, 75)=3.52, p=.065$). Lastly, participants that read a positive review admitted usually agreeing with reviews they read significantly more than participants in the negatively valenced review conditions ($F(1, 75)=4.34, p=.041$). It is debatable whether this difference originates from incomplete randomization or from a framing effect the reviews might have had.

Next, the hypotheses were tested. H1 predicted that readers of critical reviews will conform their post-viewing evaluation to the review's verdict. Both the numerical and the adjective evaluation measures were used as dependents. When comparing the two valence conditions, no significant effects were returned for either variable (numerical: $t(36)=.50$, adjectives: $t(36)=.62$). Insignificant results were also found when comparing the two conditions to the control group. On the basis of these results, H1 is rejected.

The second hypothesis predicts that critical reviews decrease measures of involvement while increasing agreement on analytical elements. Neither the positive nor the negative critical review conditions differed significantly from the respective consumer review condition (positive: $t(37)=-.27$, negative: $t(35)=1.15$). Testing for agreement on analytical elements was complicated by the uniformity of results from the experiential and critical statements from section 2 of the questionnaire. This finding meant there was no support for

using some statements as analytically framed while using the others as experientially oriented. Two out of three statements that were identified as critical ('This movie was well made' and 'this movie had good acting talent' but not 'This movie had a beautiful plot') however, differed significantly when tested for conformity to the positive or negative critical reviews; 'this movie was well made' was higher for the positive critical review condition than for the negative critical review ($t(36)=1.75$, upper tail $p=.044$), and the same result was seen for the statement 'this movie had good acting talent' ($t(36)=1.89$, upper tail $p=.034$). A comparison of the two review styles yielded a significant result for the adjective 'pathetic' - which was removed from the adjective scale during data-analysis - in the positive valence ($t(37)=2.14$, two-tailed $p=.039$). The readers of the positive consumer review rated the movie as more 'pathetic' than the participants who read the positive critical review. This result might simply be an artefact from performing t-tests on a large number of variables, increasing the odds of a type 1 error. There is no evidence that the critical reviews decrease involvement over the consumer reviews or the control group, although there is some indication that the critical review conforms viewers to critical statements about the movie. Overall, though, the second hypothesis is rejected.

The third hypothesis predicted that consumer reviews also conform their viewer to the valence of the review. This hypothesis was supported; numerical evaluation was higher for the positive than the negative consumer review conditions ($t(36)=2.80$, upper tail $p=.004$), the adjective evaluation list showed the same result ($t(36)=2.39$, upper tail $p=.011$), and an anova for both consumer valence conditions with the control group was significant (numerical: $F(2, 47)=3.84$, $p=.029$, adjectives: $F(2, 47)=3.83$, $p=.029$). When applying a Bonferroni-corrected test for the directions of differences, neither group differed significantly from the control group on the numerical evaluation. The adjective evaluation showed a difference between the negative consumer review and the positive consumer review and the control group ($t(29)=2.20$, two-tailed $p=.036$). The third hypothesis was accepted. The positive consumer review also raised answers to the involvement statements significantly ($t(36)=1.97$, upper tail $p=.029$). The critical statements that differed significantly for the critical review differ for the consumer reviews as well ('this movie was well made': $t(36)=3.80$, two-tailed $p=.001$) 'this movie had good acting talent': ($t(36)=3.23$, two-tailed $p=.003$). Contrarily to the critical reviews, one of the individual experiential statements ('Watching this movie was a beautiful experience') was also significant ($t(36)=1.92$, upper tail $p=.0315$).

H4 concerned the differences between the critical and consumer styles, and predicted that the professional nature of the critical review would increase conformity to the review valence, resulting in an interaction effect. A two-factor analysis of variance was performed with valence and style of the review. Numerical evaluations were significantly predicted by the valence of the review ($F(1, 83)=7.30, p=.008$), though not the style ($F(1, 83)=.88$). No significant interaction effect was found ($F(1, 83)=1.79$). The result is replicated with adjective evaluation as the dependent variable (main effect of valence: ($F(1, 83)=5.28, p=.024$) main effect of style: $F(1, 83)=.96$, interaction effect: $F(1, 83)=.68$). Figure 2 displays the relationship found between the treatment conditions for the numerical evaluation.

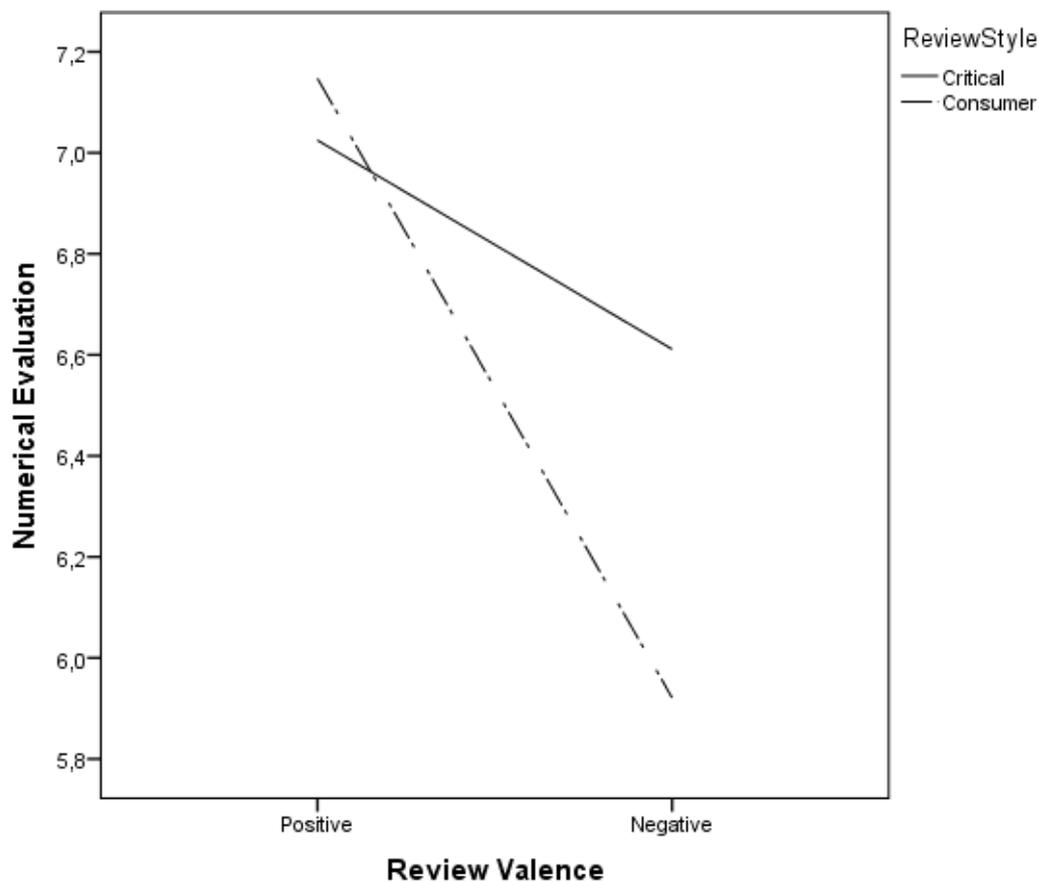


Figure 2; Plot showing the distribution of numerical evaluations across conditions. Review style is separated into critical (continuous line) and consumer (dashed line), while the review valence is shown on the horizontal axis. Control group is not shown, its average is 6.8.

While fig. 2 ostensibly indicates a difference between the review style, no main effect was found. The difference between the negative critical and consumer reviews was therefore scrutinized with an independent sample t-test. This resulted in a $t(35)$ -value of 1.66 with a two-tailed significance of .106. Since the average difference is not in the direction hypothesized (i.e. greater conformity in the critical conditions), the significance was not halved. This p -value suggested differences existed between the two review styles. On the basis of these results, H4 could not be maintained. If differences between the styles had been significant, though, they would be in the opposite direction than was hypothesized.

The final hypothesis concerned a possible overriding factor of review influence; empathic concern. When tested in a linear regression with the empathy scale as independent variable, numerical evaluation is significantly predicted ($F(1, 86)=14.50, p<.001, R^2=.38$). When the test is repeated with involvement as the dependent variable, results are again significant ($F(1, 86)=30.10, p<.001, R^2=.51$). The conceptual model from the introduction charts empathic influence on numerical evaluation through involvement. In a linear regression where involvement and trait empathy were taken together as independent variables, the direct influence of empathy on numerical evaluation (coefficient $t(86)= -.40$) disappeared, while involvement was a significant predictor (coefficient $t(86)=10.47, p<.001$). This model increased the R^2 to .79. The fifth hypothesis is accepted.

The conceptual model displayed in Fig. 1 could not be tested directly, since no direct measure of reception mode was employed. The remaining factors were used as explanatory variables of movie evaluation in wider analyses of variance. In a linear regression model with review style, valence, trait empathy, involvement and review trustworthiness - discounting the control group - as independent variables and the numerical evaluation as dependent variable, the model reached significance ($F(5, 70)=25.73, p<.001$) with an adjusted R^2 of .62. However, all individual effects apart from involvement ($t(74)=9.33, p<.001$) were insignificant. The conceptual model from the introduction pointed to involvement as a primary influencer of evaluation, mediating the influences of trait empathy and review style. The current results corroborated this contention. Since the items comprising the involvement scale can also be seen as signifiers of quality, involvement could have a confounding effect because of collinearity. These suspicions were not confirmed in an analysis of variance wherein involvement was taken as a dependent variable and numerical evaluation as an independent.

Empathy was still a significant predictor of involvement despite using numerical evaluation as an independent ($F(1, 81)=12.70, p=.001$).

For the current model to be tested as a whole, factors needed to be observed as independent as well as dependent variables concurrently. To this end, structural equation modelling was employed. The conceptual model was adapted to allow for SEM analysis; review style was set to predict involvement directly (since no direct measure for reception mode was present), and trustworthiness was influenced by style and valence. This model showed good fit with numerical evaluation as the dependent variable: $X^2(20)=7.25, p=.404, X^2/df=1.04, TLI=.995, SRMR=.087, RMSEA=.022$ (90% confidence interval from .000 to .146). The model fit did not change drastically when the adjective scale was used as the final dependent variable: $X^2(20)=7.72, p=.357, X^2/df=1.10, TLI=.985, SRMR=.089, RMSEA=.037$ (90% confidence interval from .000 to .150). Despite the good fit displayed for the full model, not all estimates were significant. The influences from trait empathy to involvement ($\beta=.46, p<.001$) and from involvement to numerical evaluation ($\beta=.77, p<.001$) were robust, though the direct influence of the review valence on numerical evaluation was only significant with an α of .10 ($\beta=.13, p=.089$). Review style was significant for the review's trustworthiness ($\beta=.27, p=.014$) but not for its influence on involvement ($p=.560$). Trustworthiness did not significantly influence the numerical evaluation ($p=.354$). When trustworthiness was removed from the model and a direct effect was observed from review valence on involvement, the fit measures indicated a beyond optimal fit (Numerical evaluation: $X^2(15)=1.21, p=.944, X^2/df=.24, TLI=1.086, SRMR=.0215, RMSEA=.000$ (90% confidence interval from .000 to .028), with adjective evaluation: $X^2(15)=1.65, p=.895, X^2/df=.33, TLI=1.069, SRMR=.0230, RMSEA=.000$ (90% confidence interval from .000 to .071). The TLI was higher than typically found in models of this kind, and the RMSEA also suggested a near-perfect fit. The close fit was likely caused by collinearity between the involvement scale and the two evaluation variables mentioned earlier. Although this did not mean the model was not viable, it caused an inflation of the fit measures. The fit was comparatively better than for the model that included trustworthiness. The revised model of review influence of enjoyment is displayed in figure 3. Review style did not significantly influence involvement in this model ($\beta=.06, p=.567$). Review valence significantly affected involvement ($\beta=.20, p=.045$) but narrowly missed significance at .05 level when tested for a direct influence on numerical evaluations ($\beta=.14, p=.051$). The squared multiple correlation for numerical evaluation was

.63, although this value was most likely inflated by the collinearity of involvement and numerical evaluations.

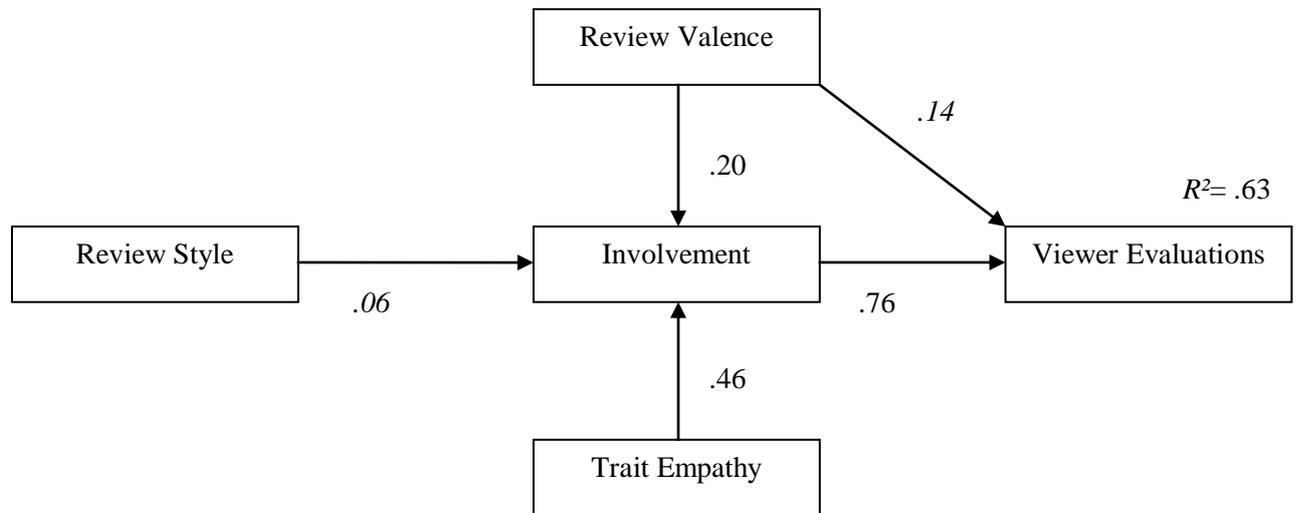


Figure 3; Model showing empirical test of the model of review influence through evaluative framing for numerical evaluation. Standardized regression weights are shown for each influence. Cursive values are not significant at .05 level.

4. Discussion

The results did not fully support the hypotheses. The two critical reviews written for this study failed to significantly steer participants' verdicts in their respective directions, although participants were aware of the valence of the review they read. The critical reviews also had seemingly no effect on the level of involvement experienced with the content as compared to the consumer reviews, though there are hints that the exact nature of the involvement did vary between review styles. The consumer reviews fared better than the critical reviews, displaying a significant difference between the positive and negative valence on the participants' final evaluations as well as their level of involvement. Though the difference between the two styles and the interaction effect with valence was not significant, the distribution of the data in table 2 and figure 2 suggests that while the positive reviews did not differ, the negative reviews diverged .7 out of 10 points with a t-test barely missing significance. There are thus strong indications that the negative consumer review polarized its readers more than the critical negative review, which is also reflected in the extreme grades the participants estimated the consumer reviewers gave the movie.

The rationale behind the fourth hypothesis - predicting more conformity in critical evaluations than consumer evaluations - was that because the participants in the critical treatments were moved towards a critical reception mode, the participants' views would be more clearly influenced by the analysis. Conversely, the experiential viewpoint offered by the consumer reviews would leave the participants to formulate their own opinions more, consequently showing less conformity than their critical treatment counterparts. Based on results, there are indications that the opposite is true, necessitating an adjustment of the theory. With the current sample, entertainment content, and fictional reviews it seems that the consumer reviews were more successful at polarizing their readers simply because they were themselves more polarised. The negative consumer review likely lowered involvement because its readers did not expect the film to be worthwhile and did not allow themselves to be involved. The review could be said to have affected both the critical and affective evaluations through the weight of its emotional plea alone, despite appearing as less trustworthy and lacking in nuance.

It is too early to deny effects of review style totally. The critical reviews were, after all, deemed more trustworthy and less polarised than the consumer reviews. It can be speculated that the perceived moderation caused the lack of effects between the two critical reviews on involvement and evaluations, though it is harder to grasp why no main effect for the review style was found. Since this study was, to our knowledge, a first exploratory foray into comparing the reviews written by professional critics with those eagerly and often posted by intrinsically motivated 'consumers' for impact on a reader's post-viewing evaluation, it would be interesting to see how the style dichotomy would fare in follow-up studies with different entertainment content.

The two positive reviews, meanwhile, showed no differences in involvement or final evaluation and demonstrated a ceiling effect. The participants from the control condition rated the movie an average of 6.8 out of 10, a mere .2 and .3 points lower than the critical and consumer positive reviews, respectively. The positive reviews did not increase post-viewing evaluation significantly from the control group. Though a direct comparison should not be drawn, the mean grade of 7.9 out of 10 the movie received on IMDB.com starkly contrasts with the ceiling of 6.8 - 7.1 seen in the current sample. Possible causes for this might be that *I'm Here* shows a selection effect among the general populace (1), attracting lovers of quirky short independent films while remaining invisible to the average individual; and that the

participants were less interested in the movie because they felt forced into watching it (2). The results corroborate those found by Basuroy et al. (2003), with larger effect sizes for negative influences. It appears to be easier to dampen a viewer's enjoyment of a movie than heighten it.

The relationship between trait empathy, involvement with the content and post-viewing evaluation was confirmed with the current data. Greater empathizers showed more involvement and consequently enjoyed the movie more than respondents scoring lower on the scale, and effect sizes indicate the theme and genre of the movie was well-suited to test this, as indicated by de Wied et al. (1994). The higher score on the involvement scale could indicate an experiential reception mode, putting any possible influence the reviews might have had into perspective. While reviews might raise or lower the final evaluation somewhat, other, individual differences obviously account for much larger variance (Eliashberg & Shugan, 1997).

The data underline the notion forwarded in the literature (e.g. Ross et al. 1975) that a source's credibility does not always exert an influence on that source's effects on attitudes and behaviour. Review trustworthiness did not significantly moderate effects exerted by the review valence. Also, despite the main effects the valence and style had on trustworthiness, only the consumer negative treatment differed significantly from the control group. Of the four reviews, the consumer negative review was seen as the least trustworthy (see fig. 2), scoring an average of 2.7 out of 7 points. The perception of this review as not trustworthy did not undermine its effect in lowering the viewers' involvement and subsequent evaluation of the movie. Moreover, the participants that read consumer reviews showed greater awareness of the true intent of the experiment. It is tentatively concluded that a perseverance effect ensured the influencability of the subjects in the consumer negative condition, with a caveat that the manipulation check could be contaminated through its place at the end of the questionnaire. However, the data could also be supported by a demand-effect, where the participants that were aware of the study's intent conformed their evaluation to what they thought were the experimenter's wishes. Obfuscation of a study's goals therefore still remains a worthwhile pursuit in this area of research.

The involvement scale used in the current study suffered from a number of issues. Firstly, though the scale was originally meant to measure affective involvement to indicate an experientially oriented reception mode with a small number of critical items, all of the items

showed good interreliability and could not be separated fully into "critical" and "experiential" factors. The involvement scale also displayed collinearity with the dependent variables that inflated values of explained variance. The choice was made to retain involvement for the model partly because the variables were not interchangeable; involvement completely mediated the effect empathy had on post-viewing evaluation. Secondly, even though the empirical evidence to separate involvement from the dependent variables was not wholly conclusive, the involvement scale was still used separately on the grounds of different item contents. Since involvement is, for most measures used in entertainment content, by nature a part of enjoyment (Vorderer et al. 2004), one can expect difficulties when trying to mediate enjoyment with involvement. It would, nevertheless, suit future researchers on the topic to be prudent in choosing a scale for the approximation of involvement.

The dual dependent variables of numerical and adjective evaluation discussed in the current article were employed to cast a wider net on the still elusive concept of evaluation. Although millions of users of review aggregation sites and this study's participants do not have to contemplate much aspects consciously and can readily distil their enjoyment and admiration of a movie into a single numerical grade, research charting the salient aspects of such a verdict is mired in individual differences and content type idiosyncrasies. Even while the two measures used for the current study displayed similar results with close significance values, differences - for example when comparing the negative consumer review to the control group - still emerged. Using either measure by itself, therefore, one would run the risk of overlooking effects or overestimating the breadth of a certain influence. The argument that using multiple dependent variables could lead to cherry-picking significant results does not hold in the face of the multitude of facets that likely make up a post-viewing evaluation. With a construct as broad as this, it cannot be expected to be measured with single items. Support for this contention comes from Zajonc (1980), who proposed an affective primacy in reasoning and preference generation. The numerical grade most viewers can readily produce is, in this view, not dependent on 'cold' cognition - in this case represented by the slightly warmer critical reception - but mostly on the affective response that precedes the cognitive elaboration. While this affective response can be guided beforehand by ruminations on an experience's worth, very few individuals take more than two seconds to produce a grade. The list of adjectives comprising the second dependent variable forced participants to gauge separate aspects of the experience, offering a more complete view of the evaluation which

likely caused the minor differences between the dependent variables to appear. The adjective items were responded to as quickly as the numerical grade, however, indicating reactions were still mostly affective.

The model used to attempt to explain influences reviews could have was moderately successful with the current dataset. The study focused on a single entertainment product, and used reviews that were crafted to differ only on a few key items. Yet, the model did show good fit with a small sample. While this is partly due to over inflation of fit statistics by the involvement measure, the model shows promise in charting mediators of influences on enjoyment. Review valence exerted most of its influence on evaluation through involvement, while trait empathy only affects evaluation through involvement. Expansion and division of the involvement scale into more specific measures better attuned to gauge reception modes would likely change results even when the current stimulus materials are used.

Summarizing, the current study not only corroborated results on review influences from Wyatt & Badger (1984), but also presented a psychologically oriented model that, if refined, could shed light on the mechanisms through which reviews could influence a viewer's post-viewing evaluation of a movie. The exploratory nature of the study necessitated a joining together of results from the otherwise disparate fields of fundamental psychology and consumer research, and was aimed at both delivering empirical evidence and starting efforts to develop a theory of how reviews impact evaluations. Since past examples of an artist's work are judged in deciding whether to see his or her newest work, a viewer's verdict should not remain outside of the scope of marketing studies. Also, the current field should be of interest to investigators looking for an application of evaluation research. Extrapolating the wealth of findings in fundamental stimulus-evaluation research to real-life settings would bridge a gap between fields and lead to a more thorough understanding of how individuals come to conclusions on the quality of artistic work.

5. References

Amblee, N. & Bui, T. (2007). Freeware Downloads: An Empirical Investigation Into the Impact of Expert and User Reviews On Demand for Digital Goods. *Americas Conference on Information Systems (AMCIS)*.

d'Astous, A. & Touil, N. (1999). Consumer Evaluations of Movies on the Basis of Critics' Judgments. *Psychology & Marketing* 16(8), 677-694

Austin, B.A. & Gordon, T.F. (1987). Movie genres: Toward a conceptualized model and standardized definitions. In B.A. Austin (Ed.), *Current research in film: Audiences, economics and law* (pp. 12-33). Norwood, NJ: Ablex

Basuroy, S., Chatterjee, S. & Ravid, S.A. (2003). How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star-Power, and Budgets. *Journal of Marketing* 67, 103-117

Boor, M. (1992). Relationships among ratings of motion pictures by viewers and six professional movie critics. *Psychological Reports* 70, 1011-1021

Burger, P. & de Jong, J. (1996). Zin en onzin van leesbaarheidsformules. *Onze taal* 65, 75-77

Burzynski, M.H. & Bayer, D.J. (1977). The effect of positive and negative prior information on motion picture appreciation. *The Journal of Social Psychology* 101, 215-218

Cohen, J. (2006). Audience Identification With Media Characters. In Bryant, J. & Vorderer, P. (Eds.). *Psychology of Entertainment* (pp. 199 - 213). New York, NY: Routledge

Davis, M.H. (1980). A Multidimensional Approach to Individual Differences in Empathy. *JSAS Catalog of Selected Documents in Psychology* 10, 85-104

Davis, M.H. (1983). Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology* 44 (1), 113-126

Dellarocas, C. & Narayan, R. (2006). A Statistical Measure of a Population's Propensity to Engage in Post-Purchase Online Word-of-Mouth. *Statistical Science* 21 (2), 277-285

Desai, K.K. & Basuroy, S. (2005). Interactive Influence of Genre Familiarity, Star Power, and Critic's Reviews in the Cultural Goods Industry: The Case of Motion Pictures. *Psychology and Marketing* 22, 203-223

Eliashberg, J., Elberse, A. & Leenders, M.A.A.M. (2006). The Motion Picture Industry: Critical Issues in Practice, Current Research, and New Research Directions. *Marketing Science* 25 (6), 638-661

Eliashberg, J. & Shugan, S.M. (1997). Film Critics: Influencers of Predictors? *Journal of Marketing* 61, 68-78

Fiedler, K., Armbruster, T., Nickel, S, Walther, E. & Asbeck, J. (1996a). Constructive biases in Social Judgment: Experiments on the Self-Verification of Question Contents. *Journal of Personality and Social Psychology* 71 (5), 861-873

Fiedler, K., Walther, E. Armbruster, T. Fay, D., Naumann, U. (1996b). Do You Really Know What You Have Seen? Intrusion Errors and Presuppositions Effects on Constructive Memory. *Journal of Experimental Social Psychology* 32, 484-511

Flesch, R. (1951). *How to Test Readability*. New Yorker, NY: Harper.

Gao, G.G., Gu, B. & Lin, M. (2006). The Dynamics of Online Consumer Reviews. *Workshop on Information Systems and Economics (WISE)*.

Gilbert, D.T., Krull, D.S., Malone, P.S. (1990). Unbelieving the Unbelievable; Some Problems in the Rejection of False Information. *Journal of Personality and Social Psychology* 59 (4), 601-613

Gemser, G., Van Oostrum, M. & Leenders, M.A.A.M. (2007). The Impact of film reviews on the box office performance of art house versus mainstream motion pictures. *Journal of Cultural Economics* 31, 43-63

Hadida, A.L. (2009). Motion picture performance: A review and research agenda. *International Journal of Management Reviews* 11 (3), 297-335

Holbert, R.L. & Stephenson, M.T. (2002). Structural equation modeling in the communication sciences, 1995-2000. *Human Communication Research* 28(4), 531-551

Holbrook, M.B. (1999). Popular Appeal versus Expert Judgments of Motion Pictures. *Journal of Consumer Research* 26, 144-155

Holbrook, M.B. (2005). The Role of Ordinary Evaluations in the Market for Popular Culture: Do Consumers Have "Good Taste"? *Marketing Letters* 16 (2), 75-86

Holbrook, M.B. & Hirschman, E.C. (1982). The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun. *Journal of Consumer Research* 9, 132-140

Horton, D. & Wohl, R. (1956). Mass Communication and Parasocial Interaction: Observations on Intimacy at a Distance. *Psychiatry* 19, 215-229

Hu, L.-T. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariances structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* 6(1), 1-55

Kaplan, A.M., Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons* 52, 59-68

Kay, J. (2009, October 26). Paranormal Activity gives spirited performance at US box office. *The Guardian*. Retrieved from <http://www.guardian.co.uk/film/filmblog/2009/oct/26/paranormal-activity>

Koehler, D.J. (1991). Explanation, Imagination, and Confidence in Judgment. *Psychological Bulletin* 110 (3), 499-519

Liebes, T. & Katz, E. (1986). Patterns of Involvement in Television Fiction: A Comparative Analysis. *European Journal of Communication* 1 (2), 151-171

Litman, B.R. & Kohl, L.S. (1989). Predicting financial success of motion pictures: The '80s experience. *Journal of Media Economics* 2 (2), 35-50

McCloskey, M. & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General* 114, 1-16

Plucker, J.A., Holden, J. & Neustadter, D. (2008). The Criterion Problem and Creativity in Film: Psychometric Characteristics of Various Measures. *Psychology of Aesthetics, Creativity and the Arts* 2 (4), 190-196

Plucker, J.A., Kaufman, J.C., Temple, J.S. & Qian, M. (2009). Do Experts and Novices Evaluate Movies the Same Way? *Psychology & Marketing* 26 (5), 470-478

Reinstein, D.A. & Snyder, C.M. (2005). The influence of expert reviews on consumer demand for experience goods: a case study of movie critics. *The Journal of Industrial Economics* 53 (3), 27-51

Ross, L., Lepper, M.R. & Hubbard, M. (1975). Perseverance in Self-perception and Social Perception; Biased Attributional Processes in the Debriefing Paradigm. *Journal of Personality and Social Psychology* 32, (5), 880-892

Shen, W. (2009). Competing for Attention in Online Reviews. *Americas Conference on Information Systems 2009 Doctoral Consortium*, Paper 2

Tourangeau, R. & Rasinski, K.A. (1988). Cognitive Processes Underlying Context Effects in Attitude Measurement. *Psychological Bulletin* 103 (3), 299-314

Tsang, A.S.L. & Prendergast, G. (2009). Is a "star" worth a thousand words?: The interplay between product-review texts and rating valences. *European Journal of Marketing* 43 (11/12), 1269-1280

Vorderer, P. (1993). Audience involvement and program loyalty. *Poetics* 22 (1 - 2), 89-98.

Vorderer, P., Klimmt, C. & Ritterfeld, U. (2004). Enjoyment: At the Heart of Media Entertainment. *Communication Theory* 14 (4), 388-408

Wang, A. (2005). The Effects of Expert and Consumer Endorsements on Audience Response. *Journal of Advertising Research* 45, 402-412

Wegner, D.M., Beattie, A.E., Wenzlaff, R. & Kerker, R.M. (1981). Incrimination Through Innuendo; Can Media Questions Become Public Answers? *Journal of Personality and Social Psychology* 40 (5), 822-832

de Wied, M, Zillmann, D. & Ordman, V. (1994). The role of empathic distress in the enjoyment of cinematic tragedy. *Poetics* 23 (1 - 2), 91-106

Wirth, W. (2006). Involvement. In Bryant, J. & Vorderer, P. (Eds.). *Psychology of Entertainment* (pp. 199 - 213). New York, NY: Routledge

Wyatt, R.O. & Badger, D.P. (1984). How Reviews Affect Film Interest and Evaluation. *Journalism Quarterly* 61 (Winter), 874-878

Zajonc, R.B. (1968). Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology* 9 (2), 1-27

Zajonc, R.B. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist* 35 (2), 151-175

Zillmann, D. (1994). Mechanisms of emotional involvement with drama. *Poetics* 23 (1 - 2), 33-51

6. Appendices

6.1 Appendix A: List of items (original)

Category/Item	Mean	Standard Deviation
Als ik deze film een cijfer van 1 tot 10 zou moeten geven, waarbij 1 het slechtst en 10 het best is, dan geef ik het een:	6.69	1.37
Evaluational Adjectives ($\alpha = .79$)	4.40	.75
Interessant	4.93	1.46
Leuk	4.65	1.52
Onderhoudend	4.49	1.34
Langdradig	4.98	1.45
Subtiel	4.32	1.36
Mooi	4.98	1.49
Poëtisch	4.86	1.63
Pretentius	4.43	1.12
Kunstzinnig	5.50	1.31
Kort	3.31	1.41
Droevig ¹	5.31	1.18
Zielig ¹	3.27	1.35
Grappig	3.14	1.52
Serius	4.80	1.37
Spannend	2.76	1.48
Involvement Statements ($\alpha = .88$)	4.41	1.09
Deze film zet je aan het denken.	4.82	1.59
Deze film is goed gemaakt.	5.02	1.44
Het zien van deze film was een mooie ervaring.	4.51	1.47
Er werd goed geacteerd in deze film.	4.47	1.54
Deze film heeft een mooi verhaal.	5.14	1.46
Ik leefde met Sheldon (de hoofdrolspeler) mee.	5.25	1.38
Ik leefde met Francesca (de vrouwelijke robot) mee.	3.58	1.44
Ik voelde me verbonden met het lot van de karakters in deze film.	4.12	1.67
Sheldon doet me denken aan mijn eigen leven.	2.76	1.72

Empathy Scale ($\alpha = .82$)	4.94	.64
Empathic Concern		
Als ik iemand zie waarvan misbruik wordt gemaakt, wil ik die persoon vaak beschermen.	5.74	.88
Als ik iemand zie waarvan misbruik wordt gemaakt, wil ik die persoon vaak beschermen.	6.02	1.12
Ik maak me vaak zorgen over mensen die slechter af zijn dan ik.	4.55	1.22
Ik zou mezelf omschrijven als iemand die snel ontroerd is.	4.72	1.54
Soms heb ik geen medelijden met mensen die problemen hebben.	4.36	1.50
Ik stoer me meestal niet aan andermans tegenspoed.	4.97	1.16
Ik zie vaak dingen gebeuren die me echt raken (Ontroeren)	4.65	1.46
Fantasy		
Als ik een interessant verhaal lees stel ik me voor hoe ik me zou voelen als de gebeurtenissen in het verhaal mij zouden overkomen.	5.42	1.41
Ik raak erg betrokken bij de gevoelens van karakters in een verhaal.	5.41	1.26
Ik kijk meestal objectief naar een film, en laat me niet meesleuren door het verhaal.	5.55	1.21
Als ik een toneelstuk of film kijk voel ik me wel eens als een van de karakters.	4.45	1.59
Vaak fantaseer ik over dingen die me zouden kunnen overkomen. ¹	5.15	1.51
Maar zelden voel ik me heel erg betrokken met een verhaal of film.	5.31	1.49
Als ik een goede film kijk heb ik er weinig moeite mee me in de schoenen van de hoofdrolspeler te plaatsen.	5.66	1.11
Perspective Taking		
Voordat ik kritiek lever op iemand, stel ik me voor hoe het is om in hun schoenen te staan.	4.49	1.36
Als ik zeker weet dat ik gelijk heb op een bepaald punt luister ik meestal niet zoveel naar andermans argumenten.	4.13	1.46
Ik probeer mijn vrienden soms beter te begrijpen door mezelf te zien vanuit hun perspectief.	4.95	1.31
Er zitten altijd twee kanten aan een verhaal, en ik wil ze altijd graag allebei horen.	5.24	1.30
Soms vind ik het moeilijk zaken uit andermans perspectief te zien.	4.28	1.53

Ik wil graag ieders kant van een onenigheid horen voor ik erover oordeel.	5.27	1.11
Als ik boos ben op iemand probeer ik me vaak in zijn of haar schoenen te plaatsen.	3.69	1.44
Review Statements (items used separately)		
De recensie was makkelijk te begrijpen.	5.74	1.25
De recensie was goed leesbaar.	5.83	1.15
De recensie was leuk om te lezen.	3.68	1.57
De recensie was betrouwbaar.	3.37	1.42
De recensent heeft goed nagedacht over zijn oordeel.	3.87	1.44
Ik lees vaak dit soort recensies voor ik een film ga zien.	2.92	1.82
Ik ben het vaak eens met de recensies die ik lees.	3.82	.91
Other Items		
Hoe heette de recensent van de recensie die je hebt gelezen?		
Wat voor cijfer denk je dat de recensent aan de film heeft gegeven (van 1 tot 10)?		
Heb je de film al eens eerder gezien (of er over gehoord)?		
Wat is volgens jou het doel van dit onderzoek?		

¹: Item was not used in further analysis.

6.2 Appendix B1: Critical positive review stimulus:

[De volgende informatie is van de website www.filmkrant.nl.]



Juli/augustus 2010, nr 253

I'm Here

***I'm Here* is een korte film geschreven en geregisseerd door de veelgeprezen Spike Jonze. In dit dertig minuten durende subtiele werk toont Jonze een diep gevoel voor de thema's die worden besproken, te weten discriminatie en toewijding.**

Robot Sheldon aan de oplader.



We volgen Sheldon, een onmiddelijk herkenbaar retro weergegeven robot die een ietwat depressief bestaan leidt als bibliothecaris. Deze androïde roept al snel dezelfde sympathie op als een *Wall-E* van Pixar als hij verliefd wordt op een onhandige vrouwelijke robot die een vrijer leven heeft, en ook het verdere verhaal geeft genoeg stof tot nadenken.

Visueel laat de film een onuitwisbare indruk achter door het gebruik van realistische effecten. De cast wordt gevuld door de verse speelstijlen van Andrew Garfield en Sienna Guillory, al blijft de invloed van deze sterretjes beperkt tot hun stemmen en mimiek.

Al met al is de lengte van deze film waarschijnlijk het enige wat hem uit de spotlights houdt. De zachte sfeer en ontroerende ontknoping maken *I'm Here* tot een onontdekte parel.

Geschreven door Jord Baucus

I'm Here

Verenigde Staten, 2010

Productie: Absolut Vodka

Regie: Spike Jonze

Scenario: Spike Jonze

Camera: Adam Kimmel

Montage: Stephen Berger, Eric Zumbrennen

Art direction: Floyd Albee

Muziek: Sam Spiegel

Met: Andrew Garfield, Sienna Guillory, Annie Hardy e.a.

Kleur, 30 minuten

Distributie: D&E Entertainment

Te zien: online

B2: Critical negative review stimulus:

[De volgende informatie is van de website www.filmkrant.nl.]



Juli/augustus 2010, nr 253

I'm Here

I'm Here is een korte film geschreven en geregisseerd door de relatief onbekende Spike Jonze. In dit dertig minuten durende weinig subtiele filmpje toont Jonze een gebrek aan gevoel voor de thema's die worden besproken, te weten discriminatie en toewijding.

Robot Sheldon aan de oplader.



We volgen Sheldon, een onrealistisch uitziende robot die een ietwat depressief bestaan leidt als bibliothecaris. Deze androïde roept helaas niet de sympathie op van een Wall-E van Pixar als hij verliefd wordt op een onbetrouwbare vrouwelijke robot die een losbandiger leven heeft, en ook het verdere verhaal lijkt het punt net te missen.

Visueel laat de film geen langdurige indruk achter door het gebruik van oninteressante computereffecten. De cast wordt gevuld door de obscure Andrew Garfield en Sienna Guillory, al blijft de invloed van deze figuren beperkt tot hun stemmen en mimiek.

Al met al is de lengte van deze film waarschijnlijk niet het enige wat hem uit de spotlights houdt. De pretentieuze sfeer en de ongeloofwaardige ontknoping maken *I'm Here* tot een protserig gebeuren.

Geschreven door Jord Baucus

I'm Here

Verenigde Staten, 2010

Productie: Absolut Vodka

Regie: Spike Jonze

Scenario: Spike Jonze

Camera: Adam Kimmel

Montage: Stephen Berger, Eric Zumbunnen

Art direction: Floyd Albee

Muziek: Sam Spiegel

Met: Andrew Garfield, Sienna Guillory, Annie Hardy e.a.

Kleur, 30 minuten

Distributie: D&E Entertainment

Te zien: online

B3: Consumer positive review stimulus:

[De volgende informatie is van de website www.moviemeter.nl.]



I'm Here (2010)



Verenigde Staten
Drama / Korte film
30 minuten

Geregisseerd door Spike Jonze
Met Andrew Garfield en Sienna Guillory

Momenteel te zien in 0 bioscopen.



Gebruikersrecensie:

Movielover14

[permalink] geplaatst op 12 augustus 2010, 20:13 uur



Ik heb net I'm here gezien en ik kan hem echt heel erg aanraden. Hij is geschreven en geregisseerd door dezelfde kerel, Spike Jonze, en die man heeft een fantastisch gevoel voor sfeer. Het is een nauw samenhangend verhaal over discriminatie en toewijding.

De film gaat over Sheldon, een schattige robot die in een bibliotheek werkt. Hij doet mij denken aan Wall-E van Pixar, en dit is net zo'n schatje. Hij komt een meisjesrobot tegen, die net zo mooi is maar dan anders vormgegeven. Zij houdt zich niet aan de regels die in deze wereld voor robots gelden, en doet waar ze zin in heeft. De rest van het verhaal verklap ik maar niet, het is veel mooier het zelf te zien. Ik vond de film mooi om te zien omdat de computereffecten goed gedaan zijn. Van geen van de hoofdrolspelers krijgen we de gezichten te zien, maar de robohoofden en mensenstemmen passen leuk bij elkaar. De acteurs, Andrew Garfield en Sienna Guillory, kende ik nog niet uit andere films, al zullen we vast snel meer van ze horen (en dan ook zien). De film is helaas maar dertig minuten, ik wenste dat hij net zo lang was als een normale film. Omdat hij behoorlijk kunstzinnig is is het misschien meer voor mensen die daarvan houden, maar ik denk dat andere mensen hem ook zeker zullen waarderen.

B4: Consumer negative review stimulus:

[De volgende informatie is van de website www.moviemeter.nl.]



I'm Here (2010)



Verenigde Staten
Drama / Korte film
30 minuten

Geregisseerd door Spike Jonze
Met Andrew Garfield en Sienna Guillory

Momenteel te zien in 0 bioscopen.



Gebruikersrecensie:

Movielover14

[permalink] geplaatst op 12 augustus 2010, 20:13 uur



Ik heb net I'm here gezien en ik kan hem echt totaal niet aanraden. Hij is geschreven en geregisseerd door dezelfde kerel, iemand die Spike Jonze heet, en die man weet in deze film helemaal niet hoe je een sfeer moet neerzetten. Het is een losjes samenhangend verhaal dat zou moeten gaan over discriminatie en toewijding. De film gaat over Sheldon, een lelijke robot die in een bibliotheek werkt. Hij doet mij denken aan een slap aftreksel van Wall-E van Pixar. Hij komt een meisjesrobot tegen, die net zo lelijk is vormgegeven (terwijl het een andere stijl is dan de eerste). Zij houdt zich niet aan de regels die blijkbaar voor robots gelden en doet waar ze zin in heeft. De rest van het verhaal verklap ik maar niet, al moet je niet teveel verwachten. Ik vond de film saai om te zien omdat de computereffecten niet speciaal zijn. Van geen van de hoofdrolspelers krijgen we de gezichten te zien, en aan het stemmenwerk te horen is dat misschien maar beter ook. De acteurs, Andrew Garfield en Sienna Guillory, kende ik nog niet uit andere films, en dat mag wat mij betreft lekker zo blijven. De film is gelukkig maar dertig minuten, en ik wenste nog steeds dat hij vroeger voorbij zou zijn. Omdat hij behoorlijk pretentius is is het misschien meer voor mensen die daarvan houden, ik denk dat andere mensen hem niet zullen waarderen.

B5 Neutral control stimulus:

[De volgende informatie is van de website www.imdb.com:]



I'm Here (2010)

29 min - [Short](#) | [Drama](#)

A young man faces the abrupt loss of his friend.

Director:

[Spike Jonze](#)

Writer:

[Spike Jonze](#)

Stars:

[Andrew Garfield](#), [Sienna Guillory](#) and [Annie Hardy](#)

Storyline

A library assistant plods through an ordinary life in LA until a chance meeting opens his eyes to the power of creativity and ultimately, love. When this new life and love begin to fall apart, he discovers he has a lot to give. *Written by [n/a](#)*

Details

Country:

[USA](#)

Language:

[English](#)

Filming Locations:

[Arlington, Texas, USA](#) [See more](#) »

Company Credits

Production Co:

[Absolut Vodka](#) [See more](#) »

Show detailed [company contact information](#) on [IMDbPro](#) »