



Case report

Based on a true story: Making people believe the unbelievable☆

Francesca Valsesia ^{*,1}, Kristin Diehl ¹, Joseph C. Nunes ¹

Marshall School of Business, The University of Southern California, 3670 Trousdale Parkway, Los Angeles, CA 90089, United States

HIGHLIGHTS

- Being labeled as “based on a true story” improves evaluation of certain stories.
- Truth-based labeling benefits stories low but not high in typicality.
- Truth based labeling heightens perceived plausibility of low typicality stories.
- Truth-based labeling does not affect perceived plausibility of typical stories.

ARTICLE INFO

Article history:

Received 11 July 2016
 Revised 27 January 2017
 Accepted 2 March 2017
 Available online xxxx

Keywords:

Story-telling
 Narratives
 Story evaluation
 Based on a true story
 Perceived plausibility
 Typicality

ABSTRACT

Storytelling is important to how people construct reality and interact with others. This research contributes to our understanding of why some stories are evaluated more positively than others, specifically how truth-based labeling (TBL), stating the story is “based on true events,” influences evaluations. Past research has failed to find an unequivocal effect of knowing a story is true on a range of responses including enjoyment, transportation, and emotional reactions. We contend this was due to past work not considering how TBL might interact with the nature of the story itself. One aspect of the story is its typicality (i.e., whether story events fall within the parameters of our past and present experiences). We propose, and show, across experimental and correlational data, that TBL increases the perceived plausibility of a story and enhances the audience's response only when a story is low in typicality to begin with. Conversely, when events in a story are already high in typicality, TBL has little effect on the perceived plausibility of the story, and in turn how the audience responds. We further provide mediational evidence for perceived plausibility as the underlying mechanism.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Human are natural storytellers. Throughout history, every society has engaged in storytelling (Brown, 1986). Stories come in various forms, from art (e.g., poems) to entertainment (e.g., novels, films), from reports (e.g., news) to word of mouth (e.g., urban legends). Storytelling is universal because it fosters social interaction and contributes to how people construe, understand, or make sense of life events, relationships, and the self (e.g., Adaval, Isbell & Wyer Jr., 2007; Baumeister, Zhang & Vohs, 2004; Chawla & Krauss, 1994). While humans tell millions of stories every day, some stories are more “successful” in terms of being evaluated more positively by others. How

successful a story is matters because the benefits of storytelling (e.g., building personal and social resources) depend on the audiences' favorable response (Reis et al., 2010).

One important aspect known to impact a story's success is its perceived plausibility, defined as whether the events in a story could occur, independent of whether or not they actually did occur (Dorr, 1983; Hall, 2003; Heath, Bell & Sternberg, 2001). Storytellers may therefore try to boost the perceived plausibility of their story, because stories low in perceived plausibility tend to be evaluated less favorably (Greenberg & Busselle, 1992). Across various forms of storytelling (e.g., books, film, oral conversations), one can observe storytellers informing their audiences that their stories are based on true or actual events, what we refer to as *truth-based labeling* (henceforth TBL). Presumably these storytellers are trying to heighten the plausibility of their stories. Note that TBL merely *suggests* that certain *elements*—not the entire story—occurred, and therefore does not apply to documentaries, newspaper articles, or biographies, in which the *entire* story is presented as true (albeit from a particular perspective).

Up until this point in time investigations into the effects of being informed a story is true have found no effects on a range of responses

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author.

E-mail addresses: valsesia@usc.edu (F. Valsesia), kdiehl@marshall.usc.edu (K. Diehl), jnunes@marshall.usc.edu (J.C. Nunes).

¹ Authors are listed in an arbitrary order and contributed equally.

including enjoyment, transportation, and emotional reactions (e.g., Ebert & Meyvis, 2014; Green & Brock, 2000). Only under certain circumscribed conditions, when highly empathetic people evaluate melodramas described as *nonfiction*, a “true story that actually happened,” have such claims been shown to affect responses to a story (Argo, Zhu & Dahl, 2008, p. 618).

Are storytellers mistaken by using TBL? We argue they are not, but that prior research has neglected to consider an important factor: the *typicality* of events within the story, and thus did not demonstrate a consistent effect of TBL on a story's success. Typicality is defined as the extent to which events within a story are commonly occurring and thus fall within the parameters of the audience's past and present, direct and indirect experiences (Hall, 2003; Shapiro & Chock, 2003). We propose, and show, that only when a story is low in typicality to begin with (i.e., includes elements inconsistent with people's past and present experiences) that TBL increases its perceived plausibility and enhances the audience's response. Conversely, when events in a story are already high in typicality, TBL is shown to have little effect on its perceived plausibility, and in turn how the audience responds.

Note that while related, perceived plausibility and typicality are distinct constructs. While it is true that a story that includes more common events should be perceived as more plausible, “plausibility does not ensure typicality” (Hall, 2003, p. 633); stories about uncommon events can be highly plausible (e.g. a friend winning the lottery). Further, other contextual aspects (e.g. source expertise) may influence plausibility independent of typicality. Relatedly, plausibility and TBL are also distinct constructs. While TBL implies some connection to reality, it is often applied liberally (Teichner, 2016) and is no guarantee audiences will consider the entire story more plausible or like it more.

We begin by showing how real world filmgoers' evaluations of movies benefit from TBL. Consider events in a story detailing how the CIA produced a fake movie in Iran to rescue American hostages, as depicted in the movie *Argo*. The events in this film are far removed from common experiences in people's lives (i.e., low in typicality). We propose exactly this and only this type of story, if labeled as based on true events, should be evaluated more favorably by the audience.

Next, a series of experiments demonstrate how the influence of TBL on evaluations depends on the *typicality* of events in the story, a critical moderator that has been overlooked in past work. We further show the effect of TBL operates through *perceived plausibility*, a belief the story *could have* happened. Labeling a story as “based on true events” heightens its perceived plausibility and improves evaluations only when it includes elements inconsistent with people's past and present experiences (i.e., low in typicality). In contrast, TBL does not have a comparable effect on the perceived plausibility of stories already high in typicality.

2. Study 1

In this study, we examine the relationship between TBL and movie evaluations in the real world by analyzing movie reviews for feature films that made the Top 100 in terms of box office revenues between 2003 and 2012, controlling for other factors. Motion pictures serve as the context for two reasons: 1) in feature films, TBL is used regularly and generously enough to allow us to assess its relationship with movie evaluations and 2) evaluations of movies are publicly available and easily attainable in the form of ratings.

2.1. Method

Our sample consists of the 1000 movies that entered the Top 100 Box Office ranking according to the website www.boxofficemojo.com in the 10-year period from 2003 to 2012 (the last year for which data was available at the time of data collection). We eliminated 98 animated movies from the dataset. Further, we omitted 18 documentaries as they

are marketed as the literal retelling of a true story. For the remaining 884 movies, we collected additional information.

The focal independent variable was whether or not the movie employed TBL (1 – yes, 0 – no). A research assistant blind to the hypotheses watched the trailer of each movie and coded whether it explicitly mentioned that the movie was either based on or inspired by true events. Overall, the dataset contains 58 movies that employed TBL (6.56%). Given 51 of 58 (88%) of the movies employing TBL were classified by the website www.rottentomatoes.com as belonging to the drama genre, our analysis focuses on the 418 movies in the dataset classified as dramas.

The focal dependent variable is *ratings*. Ratings consist of the average rating for each movie collected from the Internet Movie Database website (www.imdb.com), which aggregates evaluations from filmgoers who rate movies on a 10-point scale (1 = did not like at all, 10 = liked very much). Average ratings for dramas ranged from 3.3 to 9.0 ($M = 6.70$, median = 6.70, $SD = 0.98$).

We also collected a set of covariates to control for established drivers of movie success identified in previous literature (e.g., Elberse & Eliashberg, 2003). These include: *budget*, *MPAA rating*, *major studio*, and *star power*. *Budget* refers to the production budget for the movie, *MPAA rating* is a categorical variable capturing the suitability of the movie for different audiences, *major studio* is a dummy variable indicating whether the movie was produced by one of the “Big Six” studios, while *star power* captures the fame and audience appeal of the three main actors in the movie (for more information about these variables, see Supplementary Materials).

2.2. Results

A significant positive correlation between TBL and ratings ($r = 0.183$, $p < 0.001$) provides initial evidence that TBL has a positive influence on filmgoers' evaluations of a movie.

Next, we use multivariate regression to assess the role of TBL in influencing ratings while controlling for other known drivers of movie success. Budget information was not available for 84 movies, leaving us with a sample of 334 movies used in this analysis. Because ratings have been shown to systematically change over time (e.g., book reviews become more negative over time, Godes & Silva, 2012), we control for time dependence through a fixed effect of year by including a set of nine dummy variables. Our results are not sensitive to the inclusion of this fixed effect.

In contrast to prior research, we find that *truth-based labeling* has a significant positive effect on ratings (standardized $\beta = 0.13$, $t(318) = 2.36$, $p = 0.019$). Of the control variables, only MPAA ratings influence movie ratings. Full results are presented in the Supplementary Material. Importantly, TBL has an effect on movie evaluations over and above these controls.

2.3. Robustness checks

As a robustness check, we replicated our analysis on the full dataset, including movies of all genres and adding *genre* as a control variable. The effect of TBL is still positive and significant (standardized $\beta = 0.07$, $t(698) = 2.17$, $p = 0.030$).

Moreover, to assess the robustness of our results across different data sources of movie ratings, we collected user ratings from RottenTomatoes.com for all movies in our dataset, as well as critic ratings from Metacritic.com and estimated the same regression models as described before for these measures (see Supplementary Material). A meta-analysis assessing the average effect size of TBL on ratings across these different data sources shows a positive and significant weighted average partial correlation within dramas ($ES = 0.13$, $CI = [0.08, 0.18]$), as well as across all genres ($ES = 0.06$, $CI = [0.03, 0.10]$). All confidence intervals we report are at the 95% level. Full results are presented in the Supplementary Material.

2.4. Discussion

Using 10 years of ratings by actual filmgoers, we find that TBL is significantly and positively correlated with evaluations, providing externally valid evidence of a relationship. As with any correlational data, these findings cannot establish causation. However, these results stand in stark contrast to prior research that has attempted but failed to demonstrate any substantial effect of TBL on audiences' evaluations of a story. Note that it is possible producers apply TBL selectively to movies they fear on their own would lack typicality, films such as *Argo*. Just such a selection bias may help explain why our data exhibit an effect of TBL, whereas others did not find an effect. If only low typicality movies are promoted using TBL, the main effect observed in this study is analogous to focusing on the simple effect of TBL when typicality is low in an experiment that independently manipulates both TBL and typicality. Unfortunately, contemporaneous, independent assessments of typicality for these movies unrelated to how they were marketed are not available, making testing this assertion impossible with this data. As an alternative, Study 2 mimics such assessments using experimental evidence.

3. Study 2

In this study, we focus on a subset of movies from Study 1 that originally included truth-based labels. We asked respondents to evaluate five movies synopses. If only low typicality movies received TBL in the real world, then removing TBL should render these movies significantly less plausible in the eyes of respondents. Study 2 tests whether: 1) TBL indeed affects perceived plausibility and 2) in the absence of such labels these movies would have been seen as significantly less plausible.

3.1. Method

We recruited 894 participants through Amazon Mechanical Turk (MTurk). Here and in subsequent studies, eligible respondents were restricted to U.S. residents with a 95% or higher approval rate who had completed at least 50 hits. Participants were compensated 0.35 USD, and only one response per IP address was allowed (Goodman, Cryder & Cheema, 2013). We report all manipulations and exclusions in all studies. Additionally, we report all measures either in the paper or in the Supplementary Materials.

Our target sample was set to at least 60 respondents per cell who had not previously seen the movie. This sample size was determined based on a pre-test intended to ensure adequate power ($1-\beta > 0.80$) to detect a small effect (Cohen's $d = 0.25$). Screening of respondents was monitored using answers to the question: "Have you ever seen this movie?" No data analysis was conducted until all cells were filled (for the distribution of participants across cells, see Supplementary Materials).

The study utilized a 2 (TBL: Yes vs. No) \times 5 (Movie replicates: *Calendar Girls*, *Monster*, *Radio*, *Seabiscuit*, *The Texas Chainsaw Massacre*) between-subjects design. The five movies used in this study include all 2003 films from our database described as being based on true events. We selected movies from the first year in our sample to minimize the chance respondents today would remember whether or not the story in fact was based on actual events, allowing us to manipulate TBL effectively.

Participants read a short film synopsis adapted from www.imdb.com that omitted any information regarding the factual nature of the story's events (see Supplementary Materials). Participants were instructed that their goal was to evaluate the perceived plausibility of that movie synopsis. We manipulated TBL by varying whether or not participants were told that the story they were about to read was based on true events. After reading the synopsis, respondents were asked to rate the perceived plausibility of the story on a 9-point scale (1 = not at all plausible; 9 = extremely plausible) as well as its

plausibility relative to other dramas (-50 = Less plausible, 50 = More plausible). Finally, participants reported whether or not they had seen the movie.

We only analyze data from respondents who reported never having watched the assigned movie and who passed the manipulation check. These exclusion criteria allowed for a cleaner manipulation of TBL. Out of the 894 participants originally recruited, 271 had seen the movie they were assigned, while another 158 failed the manipulation check (final $N = 465$). The manipulation check read: "The movie you read about today was: (a) based on true events, (b) fictional, (c) you didn't tell me." Note respondents may have failed the manipulation check because they did not pay attention or because they had prior information about the movie that conflicted with the manipulation.

3.2. Results

A 2×5 between-subject ANOVA with *perceived plausibility* as the dependent variable revealed two main effects. Not surprisingly, certain movies were perceived to be more plausible than others ($F(4,455) = 50.12, p < 0.001, \omega_p^2 = 0.299$). More importantly, we find a significant main effect of TBL. Respondents who were *not* told that the movie was based on true events found the story to be *less* plausible ($M_{\text{NoTBL}} = 5.56, SD = 2.43, CI = [5.24, 5.88]$) compared to respondents who were told it was based on true events ($M_{\text{TBL}} = 6.03, SD = 2.28, CI = [5.74, 6.32]$), $F(1,455) = 6.96, p = 0.009, \omega_p^2 = 0.013$, suggesting the use of TBL when the movie was released in 2003 should have boosted its plausibility. Notably, this effect does not vary by movie (Interaction: $F(4, 455) = 0.13, p = 0.971, \omega_p^2 = -0.008$). Using perceived *relative* plausibility as the dependent variable yields similar results ($M_{\text{NoTBL}} = 3.92, SD = 24.58, CI = [0.67, 7.17]$ vs. $M_{\text{TBL}} = 8.32, SD = 24.76, CI = [5.22, 11.47]$), $F(1,455) = 5.77, p = 0.017, \omega_p^2 = 0.010$). Analyses at the movie level are provided in the Supplementary Materials.

3.3. Discussion

In an attempt at turning back time, Study 2 shows how movies promoted in real life using TBL are perceived as less plausible when these labels are absent. This finding suggests TBL is applied selectively by filmmakers to atypical stories to boost their perceived plausibility.

4. Study 3

In this study, we test our conceptual model by independently manipulating a core component of the story being evaluated—namely typicality—as well as whether TBL is employed or not. Doing so allows us to test whether the effect of TBL on evaluations of a story depends on the story's typicality as we propose.

4.1. Method

We recruited 150 participants through MTurk. Sample size was determined based on a pre-test intended to ensure that the final sample size would provide enough power ($1-\beta > 0.80$) to detect a small effect (Cohen's $d = 0.25$). Participants were compensated 0.50 USD. No exclusions were made.

The study utilized a 2 (TBL: True vs. Fiction) by 2 (Typicality: Higher vs. Lower) between-subjects design. Participants were asked to evaluate the summary of what they believed was the idea for a movie. For the focal story, we selected a relatively unknown movie that never made it to the Top 100 box office list. The movie, *Compliance*, was released in 2012, shown in only 21 theatres in the U.S., and grossed a mere \$319,285 during its theatrical release. In actuality, it is inspired by true events that took place in several U.S. restaurants between 1994 and 2004. The movie depicts an elaborate hoax involving a man identifying himself as a police officer. The alleged officer calls the fast food chain ChickWich and informs the manager that a restaurant

employee has stolen money from a customer. Further, he orders the restaurant staff to engage in outlandish behavior toward the suspect (e.g., conducting a strip search).

Details of the original story were altered to create two versions of the story that varied in typicality. The lower typicality version included a number of extreme events which were replaced by less extreme events in the higher typicality version. More specifically, in the lower typicality version, the staff follows all orders including strip searching the employee, making her do jumping jacks in the nude, and having someone spank her in an office at the back of the restaurant. In the higher typicality version, the staff plays along with the caller, but the outlandish orders are not followed. The two versions are juxtaposed in the supplementary materials. A separate test involving 305 mTurkers compared the two versions in terms of their typicality (0 = Not at All Typical, 11 = Very Typical) and confirmed the manipulation was successful ($M_{\text{Lower}} = 2.93$, $SD = 2.49$, $CI = [0.67, 7.17]$ vs. $M_{\text{Higher}} = 3.72$, $SD = 2.55$, $CI = [3.31, 4.12]$, $F(1303) = 7.34$, $p = 0.007$, $\omega_p^2 = 0.020$, see Supplementary Material for details).

Participants were randomly assigned to read either the higher or lower typicality version. We manipulated TBL by telling half the participants the story they were about to read was based on true events, while the other half was told the story was entirely fictional. After reading the story, participants rated their interest in and enjoyment of the story, as well as the likelihood they would watch a whole movie based on the story they read. The first two items are well-established measures for media content evaluations (Nabi, Finnerty, Domschke & Hull, 2006; Perry, Jenzowsky, Hester, King & Yi, 1997), while the third represents a relevant behavioral intention measure. All measures were on 9-point scales with higher values being more positive. Factor analysis reveals these three items load on one single factor ($\alpha = 0.83$) and, consequently, they were averaged to form a single evaluation index. Respondents subsequently also rated the perceived plausibility of the story using a 9-point scale (see Supplementary Materials for exact wording of all questions).

4.2. Results

A between-subjects ANOVA with *evaluation* as the dependent variable reveals only a significant interaction between TBL and typicality ($F(1146) = 4.13$, $p = 0.044$, $\omega_p^2 = 0.021$). As depicted in Fig. 1, simple

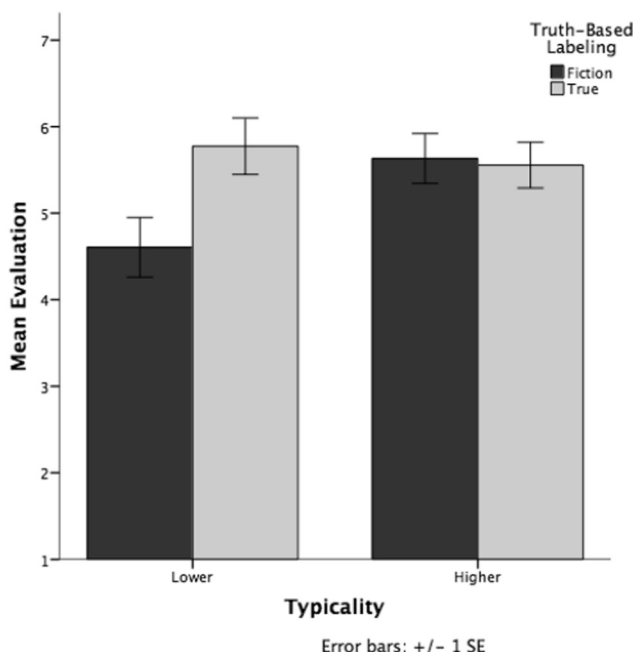


Fig. 1. The interactive effect of TBL and typicality.

contrasts indicate TBL has a positive influence on evaluations of the version lower in typicality ($M_{\text{True}} = 5.77$, $SD = 1.90$, $CI = [5.11, 6.44]$ vs. $M_{\text{Fiction}} = 4.61$, $SD = 2.12$, $CI = [3.91, 5.30]$, $F(1146) = 6.98$, $p = 0.009$, $\omega_p^2 = 0.039$), but not of the version higher in typicality ($M_{\text{True}} = 5.56$, $SD = 1.65$, $CI = [5.02, 6.09]$ vs. $M_{\text{Fiction}} = 5.63$, $SD = 1.80$, $CI = [5.05, 6.22]$, $F(1146) = 0.03$, $p = 0.857$, $\omega_p^2 = -0.007$).

Similarly, a between-subjects ANOVA with *perceived plausibility* as the dependent variable reveals only a significant interaction between TBL and typicality ($F(1146) = 4.63$, $p = 0.033$, $\omega_p^2 = 0.024$). Truth-based labeling increases the perceived plausibility of the version lower in typicality ($M_{\text{True}} = 5.53$, $SD = 2.87$, $CI = [4.53, 6.53]$ vs. $M_{\text{Fiction}} = 4.08$, $SD = 2.61$, $CI = [3.22, 4.94]$, $F(1146) = 5.71$, $p = 0.018$, $\omega_p^2 = 0.031$), but does not affect perceived plausibility of the version higher in typicality ($M_{\text{True}} = 5.85$, $SD = 2.35$, $CI = [5.09, 6.61]$ vs. $M_{\text{Fiction}} = 6.21$, $SD = 2.46$, $CI = [5.41, 7.00]$, $F(1146) = 0.38$, $p = 0.538$, $\omega_p^2 = -0.004$).

To test whether TBL improves evaluations of a story by increasing perceived plausibility and does so only when low in typicality, we estimated a moderated mediation model with 5000 bias corrected bootstrap samples (Hayes, 2013, model 7). The correlation between our proposed mediator (*plausibility*) and our dependent variable (*evaluation*) is $r = 0.522$ ($p < 0.001$). As expected, we observe that when typicality is lower, the total effect of TBL on evaluations is fully mediated by perceived plausibility ($\beta = 0.53$, $SE = 0.25$, $CI = [0.05, 1.06]$). However, this is not the case in the higher typicality version ($\beta = -0.13$, $SE = 0.20$, $CI = [-0.56, 0.24]$). The difference between the indirect effects across conditions is significantly different from zero ($\beta = -0.66$, $SE = 0.33$, $CI = [-1.36, -0.06]$), providing formal evidence of moderated mediation.

4.3. Discussion

In Study 3, we demonstrate that TBL has a differential effect on evaluations depending on the level of typicality of the story. As predicted, when lower in typicality, TBL heightens perceived plausibility and, in turn, evaluations. However, when higher in typicality, TBL conveys no observable benefit.

5. Study 4

In the last study, the two versions of the story differed in a number of details altered to manipulate typicality. In Study 4, we hold the story constant and independently manipulate story typicality.

5.1. Method

Two-hundred and forty-five undergraduates from a major West Coast university completed this study for partial course credit. Sample size was determined by the number of attendants over two standard week-long lab sessions. No exclusions were made.

The study employed a 2 (TBL: Yes vs. No) by 2 (Typicality: Higher vs. Lower) between-subjects design. Unlike the previous study, all participants in this study evaluated the exact same summary of the movie *Compliance*, which was the lower typicality version employed in Study 3. Typicality was manipulated in this study by varying whether or not respondents had been exposed to a newspaper article reporting on events similar to those in the focal story. Presumably, reading news reports depicting events similar to those in the movie should increase typicality.

First, respondents were asked to evaluate the suitability of two newspaper articles as teaching materials for non-native English speakers. For half of the participants, the first article reported on the prosecution of suspects accused of being involved in incidents similar to those depicted in the movie *Compliance*. The article was based on actual news reports of similar instances (see Supplementary Materials for exact wording). The remaining participants evaluated an unrelated

story about animal hoarding. Afterwards, all participants evaluated the same filler article concerning the U.S. economy.

Next, in an allegedly unrelated study during the same lab session, participants evaluated the same focal movie summary. Unlike in Study 3, which contrasted stories labeled as either TBL or fiction, here the story was either labeled as based on true events (TBL – Yes) or no information was provided regarding the veracity of the story (TBL – No). After reading the summary, participants completed the same dependent measures as in Study 3, namely their interest in and enjoyment of the story, as well as the likelihood they would watch a whole movie based on the story they read. Responses to these measures were highly correlated ($\alpha = 0.81$) and loaded on a single factor. They were averaged to create a composite measure of respondents' overall evaluation. Participants also rated the perceived plausibility of the story as before.

5.2. Results

A 2×2 between-subjects ANOVA with *evaluation* as the dependent variable revealed the predicted interaction between TBL and *typicality* ($F(1241) = 6.03, p = 0.015, \omega_p^2 = 0.020$) and no other significant effects. As depicted in Fig. 2, simple contrasts indicate TBL does have a positive effect on evaluation when respondents had not been exposed to similar events appearing in news stories ($M_{\text{TBL-Yes}} = 5.56, SD = 1.82, CI = [5.06, 6.05]$ vs $M_{\text{TBL-No}} = 4.74, SD = 1.85, CI = [4.29, 5.20]$, $F(1241) = 5.71, p = 0.018, \omega_p^2 = 0.019$). In contrast, TBL does not have an effect on evaluations when respondents had been exposed to stories reporting similar events ($M_{\text{TBL-Yes}} = 4.95, SD = 1.95, CI = [4.47, 5.72]$ vs $M_{\text{TBL-No}} = 5.30, SD = 1.80, CI = [4.83, 5.78]$, $F(1241) = 1.15, p = 0.285, \omega_p^2 = 0.000$).

Similarly, a between-subjects ANOVA with *perceived plausibility* as the dependent variable revealed a marginal interaction of TBL and *typicality* ($F(1241) = 3.05, p = 0.082, \omega_p^2 = 0.008$) and no other effects. Truth-based labeling does heighten perceived plausibility when participants have not been previously exposed to similar events ($M_{\text{TBL-Yes}} = 5.53, SD = 2.36, CI = [4.89, 6.17]$ vs $M_{\text{TBL-No}} = 4.57, SD = 2.35, CI = [3.99, 5.15]$, $F(1241) = 5.19, p = 0.024, \omega_p^2 = 0.017$). However, TBL does not influence perceived plausibility when participants have been previously exposed to accounts of similar events ($M_{\text{TBL-Yes}} = 4.91,$

$SD = 2.15, CI = [4.39, 5.43]$ vs $M_{\text{TBL-No}} = 4.98, SD = 2.34, CI = [4.36, 5.60]$, $F(1, 241) = 0.03, p = 0.864, \omega_p^2 = -0.004$).

Finally, just as in Study 3, we conducted a test of moderated mediation with 5000 bias corrected bootstrap samples (Hayes, 2013, model 7). The correlation between our proposed mediator (*plausibility*) and our dependent variable (*evaluation*) is $r = 0.153$ ($p = 0.017$). The results show that, while the difference between the indirect effects across conditions is not significantly different from zero at the 95% confidence level ($\beta = -0.12, SE = 0.10, CI = [-0.436, 0.004]$), in the lower typicality condition, the total effect of TBL on evaluation is fully mediated by perceived plausibility ($\beta = 0.12, SE = 0.09, CI = [0.003, 0.353]$). This is not the case in the higher typicality condition ($\beta = -0.01, SE = 0.05, CI = [-0.141, 0.088]$).

5.3. Discussion

In Study 4, we find additional evidence that the effect of TBL depends on typicality of the story. Only when typicality is relatively low does using TBL serve to heighten perceived plausibility and, in turn, significantly improve evaluations. When typicality is relatively high, no such chain of events is observed. These findings support the proposition that one reason prior research has been unable to document any consequential effect of TBL is because it did not take properties of the story itself, specifically the typicality of events within the story, into account.

6. General discussion

Storytelling is important to how people construct reality and interact with others. Telling a well-received story is critical to storytellers because it allows them to capitalize on their experiences and increase well-being (Reis et al., 2010). The present research contributes to our understanding of why some stories are evaluated more positively than others, specifically how stating the narrative is “based on true events” (TBL) influences evaluations. Past research has failed to find an unequivocal effect of knowing a story is true on a range of responses including enjoyment, transportation, and emotional reactions (e.g., Ebert & Meyvis, 2014; Green & Brock, 2000). We contend this is because past work did not consider how TBL might interact with the nature of the story itself. We show real-world movie ratings exhibit a positive relationship with TBL. By identifying a robust interaction between TBL and the typicality of the story on evaluations, our experiments help explain the discrepancy between the real world phenomena and the absence of such an effect in prior research. Truth-based labeling heightens evaluations for narratives inherently low in typicality, but does not have any effect when typicality is high. We further provide mediational evidence for perceived plausibility as the underlying mechanism. Prior research has linked greater plausibility with more favorable evaluations (Greenberg & Busselle, 1992). We show that for narratives low but not high in typicality, TBL increases perceived plausibility and, in turn, evaluations.

Using a multi-method approach, these findings extend and qualify prior research by identifying a higher level interaction which, when overlooked, obscures the true effect of TBL. Thus, our findings offer novel insights into how people evaluate stories and the role that perceived plausibility plays in their evaluations. These findings speak to a broad range of stories where narrators may use TBL. Paralleling the real world data, our experiments use stories that form the basis of movies. However, given participants read these stories rather than watched the movie, these stories could easily form the basis for many other forms of storytelling such as novels and urban legends.

While some readers might be concerned with the potential construct overlap between TBL and perceived plausibility, we would like to point out three aspects of our research that address this concern. First, if TBL and perceived plausibility were different instantiations of the same underlying construct, typicality should not moderate the relationship between the two. Second, if TBL and perceived plausibility

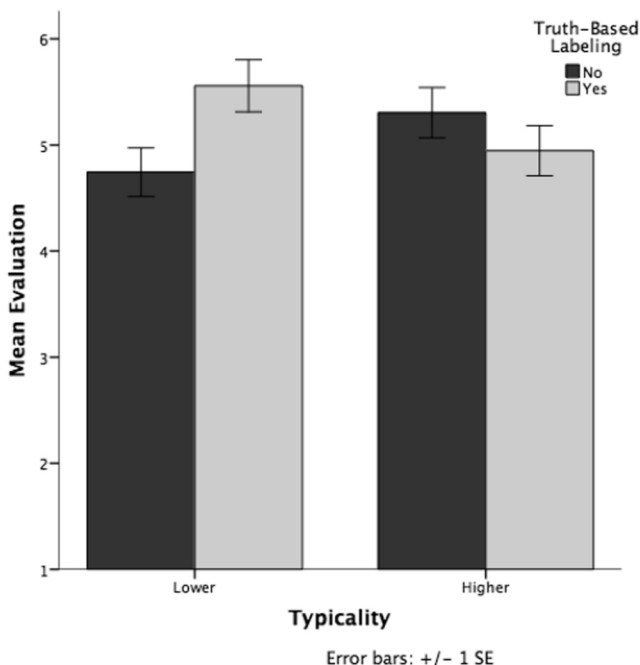


Fig. 2. The interactive effect of TBL and typicality.

were one and the same, we should not find evidence of moderated mediation whereby plausibility mediates the effect of TBL on evaluations when typicality is low but not high. Finally, we manipulate typicality independently of the narrative itself, showing that TBL can have differential effects on plausibility even when the same narrative is used as a stimulus.

While this work has focused on the effects of TBL on evaluations of a story, future research may look at when and how TBL influences how audiences evaluate the storyteller and investigate possible interactions between the storyteller, her reputation (e.g., is she known for tall tales or reporting just the facts?), and attributes of the story (such as its typicality). Such investigations may be important in understanding how people relate not just to the story but also to the storyteller and to each other.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jesp.2017.03.001>.

References

- Adaval, R., Isbell, L. M., & Wyer, R. S., Jr. (2007). The impact of pictures on narrative- and list-based impression formation: A process interference model. *Journal of Experimental Social Psychology, 43*(3), 352–364.
- Argo, J. J., Zhu, R., & Dahl, D. W. (2008). Fact or fiction: An investigation of empathy differences in response to emotional melodramatic entertainment. *Journal of Consumer Research, 34*(5), 614–623.
- Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology, 8*(2), 111–121.
- Brown, M. H. (1986). Sense making and narrative forms: Reality construction in organizations. In L. Thayer (Ed.), *Organizational communication: Emerging perspectives. Vol. 1.* (pp. 71–84). Norwood, NJ: Ablex.
- Chawla, P., & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology, 30*(6), 580–601.
- Dorr, A. (1983). No shortcuts to judging reality. In J. Bryant, & D. R. Anderson (Eds.), *Children's understanding of television: Research on attention and comprehension* (pp. 199–220). New York, NY: Academic Press.
- Ebert, J. E., & Meyvis, T. (2014). Reading fictional stories and winning delayed prizes: The surprising emotional impact of distant events. *Journal of Consumer Research, 41*(3), 794–809.
- Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science, 22*(3), 329–354.
- Godes, D., & Silva, J. C. (2012). Sequential and temporal dynamics of online opinion. *Marketing Science, 31*(3), 448–473.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*(3), 213–224.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701.
- Greenberg, B. S., & Busselle, R. (1992). Television quality from the audience perspective. *Studies of Broadcasting, 28*, 157–194.
- Hall, A. (2003). Reading realism: Audiences' evaluations of the reality of media texts. *Journal of Communication, 53*(4), 624–641.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.
- Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology, 81*(6), 1028–1041.
- Nabi, R. L., Finnerty, K., Domschke, T., & Hull, S. (2006). Does misery love company? Exploring the therapeutic effects of TV viewing on regretted experiences. *Journal of Communication, 56*(4), 689–706.
- Perry, S. D., Jenzowsky, S. A., Hester, J. B., King, C. M., & Yi, H. (1997). The influence of commercial humor on program enjoyment and evaluation. *Journalism and Mass Communication Quarterly, 74*(2), 388–399.
- Reis, H. T., Smith, S. M., Carmichael, C. L., Caprariello, P. A., Tsai, F. F., Rodrigues, A., & Maniaci, M. R. (2010). Are you happy for me? How sharing positive events with others provides personal and interpersonal benefits. *Journal of Personality and Social Psychology, 99*(2), 311.
- Shapiro, M. A., & Chock, T. M. (2003). Psychological processes in perceiving reality. *Media Psychology, 5*(2), 163–198.
- Teichner, Martha (2016). Reel life: When movies are “based on a true story” CBS news online, published February 28. Accessed December 7, 2016 at <http://www.cbsnews.com/news/reel-life-when-movies-are-based-on-a-true-story/>.