Overcoming Resistance Through Narratives: Findings from a Meta-Analytic Review

Chelsea L. Ratcliff, University of Utah*

Ye Sun, University of Utah*

*Authors contributed equally to this manuscript

Authors:

Chelsea L. Ratcliff (M.A., University of Utah) is a doctoral candidate in the Department of Communication at the University of Utah. Correspondence concerning this article should be addressed to Chelsea L. Ratcliff, Department of Communication, University of Utah, 2400 LNCO, 255 S. Central Campus Drive, Salt Lake City, Utah, 84112. Tel: 801-699-6915, email: chelsea.ratcliff.bush@utah.edu, ORCID: 0000-0002-8066-1233.

Ye Sun (Ph.D., University of Wisconsin-Madison) is an associate professor in the Department of Communication at the University of Utah. Contact information: 2400 LNCO, 255 S. Central Campus Drive, Salt Lake City, Utah, 84112. Tel: (512) 913-7887, email: ye.sun@utah.edu.

Note: This manuscript is accepted for publication at *Human Communication Research*.

Abstract

To understand mechanisms underlying narrative persuasion, a growing body of theoretical and empirical work suggests that narratives reduce audience resistance, possibly via narrative engagement processes. We performed a two-part meta-analysis to synthesize this research. Part I focused on experimental studies that directly compared narratives and nonnarratives. Based on 15 effect sizes from nine experimental studies, the overall effect size was d = -.214 (p < .001), suggesting that narratives generated less resistance than nonnarratives. Part II, a synthesis of the relationship between narrative engagement and resistance, showed the two to be negatively correlated (r = -.131, p < .001) based on 63 effect sizes from 25 studies. The heterogeneity among effect sizes in Part II was explained by narrative message characteristics. Implications of our findings and directions for future research are discussed.

Keywords: meta-analysis, narrative persuasion, narrative engagement, resistance, reactance, entertainment overcoming resistance model (EORM), transportation

Overcoming Resistance Through Narratives: Findings from a Meta-Analytic Review

Communication scholars are interested in whether narrative messages can persuade, and if so, under what circumstances and how. Nonnarrative message formats include rhetorical arguments, advice, statistical evidence, and instructions (Allen & Preiss, 1997). In contrast, in a narrative format, information or a persuasive appeal is embedded within a story featuring characters and events (Braddock & Dillard, 2016). Narratives often convey insight into primary characters' beliefs, desires, or motivational states, as well as a depiction of their actions and subsequent outcomes (Bruner, 1991). Recent meta-analyses have found that exposure to narrative messages can shape individuals' beliefs, attitudes, intentions, and behaviors (Braddock & Dillard, 2016) and that there is a persuasive advantage of narrative over nonnarrative messages (Shen, Sheer, & Li, 2015).

To understand the psychological mechanisms underlying narrative persuasion (Bilandzic & Busselle, 2013), scholars have suggested that one route to persuasion may be through quelled resistance, as narratives have the unique ability to draw audiences into their storylines and characters. Early work by Green and Brock (2000) showed that highly transported individuals were less likely to engage in critical assessment of a message, measured by identifying "false notes." The extended elaboration likelihood model (E-ELM; Slater & Rouner, 2002) highlighted the idea that absorption in a narrative suppresses counterarguing, thus rendering engaged audiences open to the persuasive subtext of a story. Going further, the entertainment overcoming resistance model (EORM; Moyer-Gusé, 2008; Moyer-Gusé & Nabi, 2010) expanded the notion of message resistance beyond just counterarguing by including reactance and perceived invulnerability as two other forms of resistance.

Moyer-Gusé and Nabi (2010) pointed out, in particular, that this theorized ability of narratives to overcome resistance remained largely untested. Following their call, empirical examinations of resistance in the context of narrative messages have cropped up, taking on this question in two ways. Some studies directly compared narrative vs. nonnarrative messages in the amount of message resistance aroused (e.g., de Graaf et al., 2017), and other studies focused on aspects of narrative processing—namely, forms of narrative engagement, such as absorption in the story or merging with the character—as potential psychological correlates with various forms of resistance (e.g., Quintero Johnson & Sangalang, 2017; Reinhart & Anker, 2012).

This paper reports a meta-analysis that synthesizes findings from these experimental (Part I) and correlational (Part II) studies. The primary goal is to provide a systematic answer to the question they collectively sought to address: To what extent, if any, do narratives overcome resistance? Knowledge about whether this effect is real, and how big it is, can provide important guidance for future research (Borenstein, Hedges, Higgins, & Rothstein, 2009). Second, given the methodological differences in primary research, what variations may amplify or constrain a narrative's capacity to reduce resistance? In particular, researchers have examined the broad idea of "resistance" in its various manifestations and have used different concepts to capture narrative engagement. Studies have also investigated variations in message features of narrative stimuli. A second goal of our meta-analytic syntheses, therefore, is to identify boundary conditions, if any, for the relationship between narrative messaging and resistance.

Narratives Overcoming Resistance

According to cognitive scientist Jerome Bruner (1986, p. 11), humans use two distinct, natural types of communication to persuade: "A good story and a well-formed argument." While the latter presents supporting evidence, the former offers an example by way of a depiction of

characters and events (Braddock & Dillard, 2016). As a message strategy, narratives are increasingly used as an alternative to argumentation (Bilandzic & Busselle, 2013) and are employed in health promotion campaigns (Hinyard & Kreuter, 2007), advertising (Escalas, 2004), and educational entertainment (Moyer-Gusé, 2008), among other contexts. A few meta-analyses have shown small to moderate effects for narratives' persuasive potency. For example, narrative messages, when compared to a control group, had an overall effect of $r = .063$ ($p < .01$) on all persuasive outcomes combined (attitudes, behavioral intentions, and behaviors; Shen et al., 2015). Tukachinsky and Tokunaga's (2013) meta-analysis on the relationship between narrative engagement and persuasion showed an overall effect size of $r = .27$ ($p < .001$) across types of narrative engagement (e.g., transportation, identification, etc.) and persuasive outcomes (e.g., knowledge, attitudes, and behaviors).

In explaining *how* narratives persuade, the focus of much attention has been on "the antithesis of persuasion": resistance (Knowles & Linn, 2004, p. 3). As attempts to persuade are often met with resistance, especially when these attempts are obvious (Brehm & Brehm, 1981), a narrative message may overcome this barrier with its "unobtrusiveness of persuasive subtext" (Slater & Rouner, 2002, p. 178) and its power to engross the audience (Dal Cin, Zanna, & Fong, 2004). Audiences of narratives, often described in the metaphor of a traveler who leaves the "world of origin" and enters the "narrative world" (Green & Brock, 2000; van Laer, Ruyter, Visconti, & Wetzels, 2014, p. 799), become part of the story and experience the event vicariously with the characters. Upon exit, they emerge from the transported experience somewhat transformed, "different from the person one was before entering the milieu of the narrative" (Green, Brock, & Kaufman, 2004, p. 315). When a persuasive message is embedded

in the story and/or carried by the characters, persuasion occurs to the immersed, less critical, and less defensive "travelers."

The idea that narratives reduce resistance was proposed in transportation theory (Green & Brock, 2000, 2002) and E-ELM (Slater & Rouner, 2002), and was further developed in EORM (Moyer-Gusé, 2008, Moyer-Gusé & Nabi, 2010). In the past decade, a growing body of studies has included measures of resistance in the research on narrative persuasion. One group of experimental studies assessed the relative effect of narratives, compared to nonnarratives, as an alternative form of presenting the persuasive information and advocacy. Other studies focused on the processing of narratives—involvement with the storyline or the characters—and how such narrative engagement correlates with resistance. Our first goal is to empirically synthesize these findings to assess the ability of narratives to reduce resistance as proposed and articulated in the previous theoretical frameworks. We examine the overall effect through the following two hypotheses, on experimental data and on correlational data separately:

H1: Narrative message formats produce less resistance to persuasion than nonnarrative formats (Part I, experimental data).

H2: Higher narrative engagement is associated with lower resistance (Part II, correlational data).

Resistance and Narrative Engagement: Different Forms and Processes

Whereas resistance and narrative engagement are the theoretical cornerstones of EORM, they are also broad constructs, each subsuming different concepts that in turn are variedly the focus of different primary studies. In this section, we summarize and describe these differences.

Forms of Resistance

First, a survey of the existing literature compels a conceptual clarification of the notion of "resistance." Resistance, intuitively understood as a counterproductive force hindering persuasion, is rarely explicitly defined. In proposing EORM, Moyer-Gusé (2008) described resistance as "a reaction against change in response to some perceived pressure for change" (p. 414). In line with that, we define resistance as an individual's motivated response, triggered by the perceived persuasive attempt and enacted to disregard the intent and/or the content of persuasion. We emphasize that resistance is a reaction *during* message processing, preceding the persuasive outcome such as message acceptance or perceived effectiveness, or other attitudinal or behavioral outcomes.

In the original formulation of EORM, resistance encompasses three forms, including counterarguing, reactance, and perceived invulnerability to a health risk (Moyer-Gusé & Nabi, 2010). Counterarguing, the form of resistance of focus in E-ELM, refers to the "generation of thoughts that dispute or are inconsistent with the persuasive argument" (Slater & Rouner, 2002, p.180). Reactance (Brehm & Brehm, 1981; Dillard & Shen, 2005) is a psychological reaction, consisting of anger and negative cognition, in response to persuasive effort as a threat to freedom. Whereas narrative messages discourage counterarguing by reducing the motivation and/or abilities of audience members to engage in message scrutiny, narratives "may side step reactance" by masking the persuasive intent (Moyer-Gusé, 2008, p. 415). In our syntheses, we exclude EORM's third category, perceived invulnerability to a health risk, for two reasons. First, it is specifically about health risks, not generalizable to other persuasive situations. Second, conceptually, perceived invulnerability is more akin to risk perception, which is typically regarded as a persuasion outcome rather than an aspect of message processing (Shen et al., 2017). We therefore focus on the literature that operationalizes resistance as counterarguing or

reactance. The extant operationalizations of resistance in narrative research can be further delineated into four categories: *counterarguing*, *perceived freedom threat*, *anger*, and *message derogation*. We will briefly describe how each has been used in the primary studies.

Counterarguing. The central tenet of E-ELM is that entertainment-education programs facilitate persuasion by reducing counterarguing, which results from message scrutiny or critical message processing (Slater & Rouner, 2002). EORM further emphasizes this point, that counterarguing involves "careful attention and thoughtful elaboration on a message" under the condition that individuals are "sufficiently motivated and able to do so" (Moyer-Gusé & Nabi, 2010, p. 30). So situated, counterarguing is a concept that captures in-depth message processing in a critical manner. Narratives reduce this form of resistance by disabling conditions for systematic processing.

In narrative persuasion studies, counterarguing is operationalized as the degree to which audiences generated arguments against specific points in the message. Most studies adopted the counterarguing scale developed by Nabi, Moyer-Gusé, and Byrne (2007), including items such as "I found myself actively disagreeing with the author" and "I was looking for flaws in the author's arguments." A few studies used open-ended responses to code for refutation of specific points in a message. For example, Niederdeppe and colleagues coded specifically for refutations of external causes or society attributions of obesity as advocated in the message (e.g., responses such as "lack of sidewalks is not an excuse for not walking"; Niederdeppe et al., 2011; Niederdeppe, Roh, & Shapiro, 2015).[1]

Perceived threat to freedom. Threat to freedom is the perception that one's freedom to think, feel, or act of their own accord is threatened by the persuasive message (Brehm & Brehm, 1981; Dillard & Shen, 2005). In the psychological reactance literature, perceived threat to

freedom is strictly defined as the antecedent of the latent construct *reactance* (Dillard & Shen, 2005). In the narrative literature, however, perceived freedom threat has sometimes been used as a direct representation of reactance itself. For example, authors used threat to freedom measures to represent "cognitive reactance" (Moyer-Gusé & Nabi, 2010), or used the term interchangeably with "psychological reactance" (Reinhart & Anker, 2012). Most of the studies we have located measured freedom threat using Dillard and Shen's 2005 scale (e.g., "The message threatened my freedom to choose"). A few studies used items adapted from Lindsey's 2005 scale of reactance (e.g., Sangalang, 2015: "I am uncomfortable being told how to feel about tobacco use").

Anger. Anger is typically treated as the affective component of reactance in this literature. A persuasive message can arouse this emotional reaction as audience members resist being told what to do. Most studies used the anger index used by Dillard and Shen (2005) and developed by Dillard and Peck (2001), which consists of four items ("angry," "irritated," "annoyed," and "aggravated"). Some studies used a different measure, including anger toward the story's protagonist (Niederdeppe et al., 2014), and a scale of feeling repulsed or insulted (Krakow et al., 2018, using the scale from Madden, Allen & Twible, 1988).

Message derogation. Another category of cognitive resistance measures in the narrative literature captures negative responses that do not necessarily involve refutation of specific points in the message. Differentiable from counterarguing, such measures operationalize resistance as dismissive or derogatory reactions during message processing. More specifically, they could manifest as dismissal of the overall topic or issue in a message or criticism of the quality of the delivery (Keer, van den Putte, de Wit, & Neijens, 2013; Kim & Niederdeppe, 2016), or hostile reaction toward the story's characters or events (e.g., "I hate the character," "the event is disgusting," Zhou & Shapiro, 2017, p. 1301). Message derogation is sometimes measured as

negative responses coded from thought-listing (e.g., Shen et al., 2017) following the reactance study by Dillard and Shen (2005). Some studies used Witte's (1994) "message minimization" scale, which was derived from Witte's (1992) content analysis of negative cognitive responses to persuasion, and which asks whether audiences found a story or message distorted, overstated, exaggerated, etc. (e.g., Kim & Niederdeppe, 2016).

Processes of Narrative Engagement

Narrative engagement or involvement has been key to understanding narrative effects (Tukachinsky & Tokunaga, 2013). Between the "world of origin" and the "narrative world" (van Laer et al., 2014, p.799), audiences switch between two reception modes: an *involved* mode and a *distanced* mode (Vorderer & Hartmann, 2009). In the distanced mode, users process the media content with an analytical lens. In the involved mode, immersion cultivates "a perceptual illusion of nonmediation" (Lombard, Reich, Grabe, Bracken, & Ditton, 2000, p. 77), rendering the narrative content the reference frame for sense-making. Narrative engagement encompasses several related processes, all highlighting the convergence between the audience and the narrative world. Three commonly measured engagement variables are *transportation* into the story, empathic *identification* with characters in the story, and *parasocial interaction* (PSI) or a feeling of companionship with the characters (Tukachinsky & Tokunaga, 2013), as briefly reviewed below.[2]

Transportation. Transportation refers to the process of becoming absorbed in the plot of a story (Green & Brock, 2000). It is "an engrossing temporary experience" (van Laer et al., 2014, p. 800) that distracts audiences from critical processing or message scrutiny (Green & Brock, 2000; Moyer-Gusé & Nabi, 2010; Slater & Rouner, 2002). Transported individuals, having suspended their real-life beliefs and knowledge and compromised their motivation and ability to

evaluate the message, are more prone to persuasion. All of the studies included in our meta-analysis used Green and Brock's (2000) scale of transportation, sometimes with adaptation (e.g., Walter, Murphy, Frank, & Baezconde-Garbanati, 2017).[3]

Character identification. As articulated by Cohen (2001), character identification is another concept representing audiences' immersive experiences. Also referred to as "empathic identification" (Tukachinsky & Tokunaga, 2013), identification takes place when users merge with the character(s) in the story and have a temporary suspension of their self-concept. As they "become the character" (Cohen, 2006), users adopt the character's perspectives and feelings and in turn align their viewpoints or behaviors with those suggested in the story. Most studies used the scale of character identification from Cohen (2001) or its adapted version from Tal-Or and Cohen (2010), with the exception of one study using a measure of empathy toward story characters (Niederdeppe et al., 2014).

Parasocial interaction. Parasocial interaction occurs when an individual engages in a "quasi-social, one-direction interaction" or relationship with a character in a story (Tukachinsky & Tokunaga, 2013, p. 289). Parasocial interaction with a story's main character has been found to reduce resistance, so long as viewers did not perceive persuasive intent in the message (Moyer-Gusé & Nabi, 2010). Parasocial interaction was measured using either the scale by Rubin and Perse (1987; e.g., in Moyer-Gusé & Nabi, 2010) or that by Schramm and Hartmann (2008; e.g., Shen et al., 2017).

In the above review, we described how resistance and narrative engagement have been approached differently in primary studies. How such differences may influence the effect size of narratives overcoming resistance is an empirical question to be addressed. We therefore propose

the following research question about the potential moderating roles of different forms of resistance and narrative engagement processes:

RQ1: Does the relationship between narrative engagement and resistance vary depending on (a) form of resistance and/or (b) process of narrative engagement?

Narrative Message Features and Other Characteristics

Narrative message features have been an understudied aspect of narrative research. Previous research has examined the medium through which narratives are delivered, including textual vs. audiovisual formats (Walter et al., 2017). In addition to *medium*, narrative message stimuli used in primary studies also vary in other aspects, such as *genre* (e.g., an educational-entertainment program vs. an advocacy message), *length* (e.g., a 30-second PSA vs. a 45-minute TV show episode), and *the number of primary characters* in the narrative (single vs. multiple). We chose to focus on these four characteristics as our close reading of the literature showed them to represent a range of differences among the narrative stimuli used in primary studies. They also potentially speak to theoretical questions about the role of explicitness of persuasive intent (e.g., genre) as well as practical questions regarding message design features that may matter (e.g., length, number of primary characters).

We also investigate other methodological factors that are typically examined in meta-analyses (e.g., Chan, Jones, Hall Jamieson, & Albarracín, 2017; Dopp, Borduin, White, Mark, & Kuppens, 2017; Tukachinsky & Tokunaga, 2013). These methodological variations include demographic characteristics of the sample (e.g., average age) as well as study characteristics (e.g., the setting of data collection, publication status).

We examine the potential effects of three sets of moderators through the following research question:

RQ2: Is the relationship between narrative engagement and resistance moderated by (a) narrative message characteristics, (b) sample characteristics, and (c) study characteristics?

Methods

Literature Search

Using a Boolean search term[4], the following databases were searched with a cutoff time point of Sept. 22, 2017[5]: *PsycARTICLES*, *PsycINFO*, *Communication and Mass Media Complete*, *Academic Search Premier*, *ERIC*, *MEDLINE*, and *Psychology and Behavioral Sciences Collection*. The search included published studies, conference papers, theses, and dissertations. This search rendered an initial pool of 173 papers, from which 124 papers were removed (109 were irrelevant, nine were duplicates, and six did not contain empirical data) and 49 papers were retained. Review articles and reference lists were then scanned to ensure that no relevant studies were missed; this rendered one additional study. Lastly, because inclusion of grey literature helps reduce the impact of publication bias in meta-analyses (Borenstein et al., 2009), we contacted scholars in this area to locate unpublished papers. Ten papers were provided that met the search criteria, rendering a final pool of 60 papers for further screening.

Inclusion Criteria

These 60 papers were scrutinized to determine whether they met the inclusion criteria for Part I (experimental data) or Part II (correlational data). For Part I, we included studies that (a) used an experiment to test a narrative condition against a nonnarrative condition, where the nonnarrative was a topically relevant message[6]; and (b) measured resistance as an outcome variable following message exposure (operationalized as *counterarguing*, *anger*, *threat to freedom*, or *message derogation*). For Part II, studies were included if they measured the correlation between resistance (*counterarguing*, *anger*, *threat to freedom*, and/or *message*

derogation) and narrative engagement (*transportation*, *identification*, and/or *parasocial interaction*). Most studies in Part II used only narratives as message stimuli (with the exception of Keer et al., 2013, Kim & Niederdeppe, 2016, and Moyer-Gusé & Nabi, 2010). For the papers that did not provide the primary data needed for effect size calculation, we requested data from the authors.[7] We had to exclude two papers whose authors did not respond to our requests. After the screening, we retained a final pool of eight papers for Part I and 21 papers for Part II (see Tables 1a and 1b).

Effect Size Extraction

In this paper, the term "study" refers to each independent sample from which the effect size estimates are extracted. For example, if a paper reports data from two independent samples, we count them as two studies even though they may not be labeled as such in the authors' report. Our unit of analysis is each *conceptually distinct* effect estimate. An effect estimate is conceptually distinct if it represents a different type of narrative engagement or type of resistance. Therefore, multiple effect sizes can be extracted from a study if each captures a unique conceptual category. Statistical dependency among effect sizes from the same study was handled by using multilevel models. This decision allows us to not only capture a comprehensive picture by encompassing the different ways narrative engagement and resistance have been measured in primary studies, but also to empirically test whether such differences matter.

Effect Size Calculation

In Part I, we used Cohen's *d*, the standardized mean difference, to represent the effect of narrative vs. nonnarrative messages (Cohen, 1988).[8] A negative *d* represents *less* resistance after exposure to a narrative message compared to a nonnarrative message, and in other words a *stronger* effect of narrative in overcoming resistance.

In Part II, the sample zero-order correlation between narrative engagement and resistance was first transformed to Fisher's *z*, which was the metric used in analyses; the summary value of Fisher's *z* was then converted back into *r* for presentation and interpretation purposes (Borenstein et al., 2009). A negative Fisher's *z* or *r* means *less* resistance associated with *greater* narrative engagement.

Coding of Moderators

The coding scheme and the number of effect sizes associated with each variable or subgroup are reported in Table 2. To establish intercoder reliability, a subset of randomly-selected studies that included 15 to 20 effect sizes per variable (representing roughly 25% of the total number of effect sizes) was first coded by two independent coders. Krippendorff's alphas for these variables (Hayes & Krippendorff, 2007) ranged from .78 to 1.00. Disagreements were discussed and resolved. One author then completed coding for the rest of the data. Because of the small number of studies in Part I we did not proceed with moderator analyses due to concerns with unstable estimates. Thus, moderator analyses were performed only for Part II.

Analytic Procedures

Meta-analytic models. The choice of fixed-effect (FE) or random-effects (RE) models has been much discussed among meta-analysts (Hedges & Vevea, 1998; Hunter & Schmidt, 2000; Schmidt, Oh, & Hayes, 2009). The appropriateness of one or the other depends on the nature of the data and the researchers' assumptions. FE models, assuming one "true" effect size across all studies, are more appropriate for a pool of studies deemed largely identical in research procedures, researchers, instruments, populations from which the samples were drawn, and so on. RE models assume that there are varied "true" effect sizes across studies, and empirically assess the variance among "true" effect sizes in addition to the sampling variance. We deem the

random-effects model to be more appropriate for our data. Our conclusions will be drawn from the RE model, but results from the FE model will be presented for interested readers.

To control for the dependency among effect sizes from the same study, we performed three-level meta-analysis, traditionally referred to as "multilevel meta-analysis" (Hox, 2002; Raudenbush & Bryk, 2002). In a recent effort to clarify the terminology, Pastor and Lazowski (2018) argue that the term "multilevel meta-analysis" is redundant and can lead to misconceptions, because all meta-analytic models have at least two levels, with persons (Level 1) nested under effect sizes (Level 2). Following their suggestion, we used the term "three-level meta-analysis" to refer to data where effect sizes are further nested under studies and differentiate it from a "two-level meta-analysis" (i.e, a traditional meta-analysis assuming independence among effect sizes). A three-level random-effects model estimates the variance components at the level of effect sizes (i.e., the within-study variance, Level 2) and at the level of studies (i.e., the between-study variance, Level 3).

Assessing and explaining heterogeneity. Three diagnostic indices of heterogeneity were obtained to assess the presence of non-sampling variance (Borenstein et al., 2009). The Q test yields evidence of statistical significance of heterogeneity (i.e., whether true effect sizes differ). I^2 is the ratio of heterogeneity to the total variance. In random-effects models, T^2 or T represents the amount of "true" variability on the same metric scale as the effect size. In the presence of heterogeneity, three-level mixed-effects meta-regression models were then analyzed to identify potential moderators. These models estimated the fixed effects of moderator variables as well as the reduction in study-level and effect-size-level variances.

In the following section, we report results from Parts I and II by first presenting results from models without moderators: the mean effect sizes and evidence of heterogeneity. Then we

report results from moderator analyses using Part II data. All statistical analyses were conducted using the "*metafor*" package (version 2.0-0; Viechtbauer, 2010) in R (version 3.5.0; R Core Team, 2018) using the Maximum Likelihood (ML) estimation.

Results

Description of Studies

Most studies in both Part I and Part II were in the domain of health communication, covering topics such as binge drinking (Keer et al., 2013; Kim & Niederdeppe, 2016; Zhou & Shapiro, 2017), cigarette or drug use (Banerjee & Greene, 2012; de Graaf et al., 2017; Sangalang, 2015), sexual health (Moyer-Gusé et al., 2011; Quintero Johnson & Sangalang, 2017; Shen et al., 2017), and cancer screening (Krakow et al., 2017; Kreuter et al., 2010; Walter et al., 2017). A few papers were situated in the context of social issues such as bullying (Shade, 2017) and sexual diversity (Igartua & Vega Casanova, 2016), while others focused on support for health policy (e.g., Niederdeppe et al., 2012) and one paper investigated commercial advertisements (Krakow et al., 2018). The samples included in both analyses featured a higher percentage of female participants (M = 65.41%, SD = 17.05%, for Part I; M = 68.95%, SD = 18.84% for Part II) and White participants (M = 67.93%, SD = 31.13% for Part I; M = 56.75%, SD = 29.71% for Part II). The average sample age was 37.34 years (SD = 16.49) for Part I, and 25.48 years (SD =11.08) for Part II.

Overall Effect Size and Heterogeneity

Table 3 reports the weighted mean effect sizes for both analyses. Findings from the three-level RE models are discussed here. For Part I, based on 15 effect sizes (total N = 5,397), the overall effect size was d = -.214, p < .001, 95% CI [-.329, -.099] (the equivalent r = -.106). H1 was supported: Exposure to narrative messages, compared to nonnarrative messages, generated

less resistance. For Part II (k = 63; total N = 17,184), the weighted mean of correlations between narrative engagement and resistance was also significant: Fisher's z = -.132, r = -.131, $p < .001$, 95% CI for *r:* [-.206, -.055] (the equivalent d = -.264). H2, predicting less resistance associated with greater narrative engagement, was supported. Both were "small" effect sizes according to Cohen (1992). For Part II data, we also obtained mean effect sizes for each resistance type and narrative engagement type from the three-level RE models (see Table 4). For different types of resistance, the relationship was significant for all but anger. For narrative engagement types, the effect size was significant for transportation and identification but not for parasocial interaction.

The right section of Table 3 reports indices of heterogeneity. There was indication of heterogeneity for both analyses. In Part I, there was a significant Q test: $Q(14) = 28.69$, $p = .012$, and about 72% of the total variability was unrelated to sampling (I^2). The estimated variance of true effect sizes was .019 (T^2). For Part II, the Q test was also significant, $Q(62) = 659.80$, $p < .001$, and the I^2 statistics indicated that over 90% of the total variability was attributable to heterogeneity. The amount of "true" variance was estimated to be .043 (T^2).

Moderator Analyses for Part II

As mentioned before, the small sample size of Part I hindered us from further moderator analyses. For Part II, three-level mixed-effects meta-regressions were conducted. We fitted separate models on each set of moderator variables to avoid inflation of Type II error rates (Higgins & Green, 2011; Raudenbush & Bryk, 2002) while controlling for effects of similar covariates. For the demographic variables (sample age, female percentage, and White percentage), due to the uneven presence of missing data (i.e., certain sample demographics were not reported in primary studies), they were each examined separately to avoid substantial data loss.[9] Our moderator analyses thus involved the following models to address RQ1 and RQ2:

Models 1a through 1c (sample mean age, percentage of females, and percentage of White participants, respectively), Model 2 (study or report characteristics), Model 3 (narrative message characteristics), and Model 4 (construct operationalization variables).

Unstandardized coefficients from the meta-regression models are reported in Table 5. Given the coding of effect size (i.e., a negative *r* means less resistance associated with greater narrative engagement), a negative regression coefficient indicates that the predictor variable *enhances* the effect of narrative engagement in resistance reduction. Table 5 also provides model-level statistics for Models 2 through 4: (1) $Q_{M(df)}$ represents the moderating effect of the group of variables in the model; (2) R^2_{META} is the proportion of *true* variance explained by the moderators, reported separately for the effect size level ($R^2_{META(2)}$) and the study level ($R^2_{META(3)}$).

Results from Models 1a – 1c showed significant effects of the average sample age (b = -.008, p = .030, 95% CI: [-.015, -.001]) and percentage of female participants (b = -.524, p = .001, 95% CI: [-.845, -.202]). Studies with an older sample, or one that had a higher percentage of females, tended to observe a stronger relationship between narrative engagement and resistance. Percentage of White participants showed no significant effect.

In Model 2, the only significant effect was from the comparison between data collection in a "natural" setting vs. in a lab (b = -.294, p = .041, 95% CI: [-.576, -.013]), suggesting a stronger negative relationship between narrative engagement and resistance in the natural setting. Neither the study population nor publication status showed any significant effect.

In Model 3, narrative message characteristics were all significant predictors. Stronger effects were from messages that were of medium length, compared to short messages (b = -.408, p < .001, 95% CI: [-.631, -.185]). The difference between long vs. short messages was not statistically significant. The contrast between medium and long messages was also significant

(not shown in the table): b = -.284, *s.e.* = .087, p = .001. Education-entertainment messages produced a stronger negative relationship than messages with a clear persuasive advocacy (b = -.350, p < .001, 95% CI: [-.550, -.149]). Narratives that featured multiple main characters showed a weaker relationship than those with a single identifiable character (b = .406, p < .001, 95% CI: [.227, .584]). Message medium showed a significant effect such that textual messages produced a stronger relationship than audiovisual messages (b = -.222, p < .001, 95% CI: [-.323, -.120]).

Model 4 examined different operationalizations of narrative engagement and resistance by including both the main effect and interaction effect terms. Given that there were only five effect sizes for the category "parasocial interaction," we combined it with character identification as "character-based engagement,"[10] to compare with transportation as "content-based engagement." One interaction term was significant: The differential effect of freedom threat vs. anger was contingent upon type of narrative engagement (b = .264, p = .022, 95% CI: [.038, .489]). As depicted in Figure 1, character-based engagement had a stronger negative relationship with anger than content-based engagement; the reverse was true for freedom threat, where content-based engagement exhibited a stronger relationship than character-based engagement. There was no main effect for engagement type. For resistance measures, the difference was between the cognitive resistance measures (perceived freedom threat, counterarguing, message derogation) and the affective measure, anger (p = .037, .049, and .007, for freedom threat, counterarguing, and message derogation respectively, Table 5); there were no differences among the former three (results not displayed in the table). Greater narrative engagement was more effective in reducing cognitive forms of resistance than anger.

Model-level statistics. Model-level statistics in Table 5 showed that the group of narrative message characteristics significantly reduced heterogeneity, $Q_M(5) = 42.01$, p < .001,

and explained 100% of the true variance at the study level. Study/report characteristics as a group did not have a significant moderating effect, $Q_M(5) = 5.17$, *n.s.* Construct operationalizations were not significant moderators as a group either, $Q_M(7) = 11.49$, *n.s.*, explaining 27% of the variance at the effect size level.

Publication Bias

Publication bias, "the selective publication of studies with a statistically significant outcome" (van Assen, van Aert, & Wicherts, 2015, p. 293), is a threat to the validity of meta-analytic findings. Moderator analyses in Part II showed that publication bias was not a significant predictor. We assessed the possibility of publication bias using three established methods in the meta-analysis research.[11] First, we inspected the funnel plot, with standard errors plotted against effect sizes. Asymmetry of the plot would indicate the potential presence of bias. Second, we conducted Egger's regression test to formally test the association between effect size and standard error. Because it tends to have low statistical power (Sterne & Egger, 2006), a significance level of .10 is typically applied. Third, we performed the *Trim and Fill* procedure to identify and correct for the potential bias, which provides an adjusted estimate of the overall effect size if bias is detected (Duval & Tweedie, 2000).[12] According to Borenstein et al. (2009), a publication bias analysis serves to categorize the results into one of three scenarios: (a) trivial impact of bias, (b) impact of bias is not trivial but does not invalidate major findings, or (c) impact of bias is substantial and calls major findings into question.

As these methods were developed for traditional two-level meta-analyses, we obtained the study-level mean effect sizes and used them for publication bias analyses. For the two-level data, the estimated overall effect sizes were comparable to those from the three-level data presented earlier: In Part I, based on nine effect sizes, $d = -.207$, *s.e.* = .058, $p < 0.001$, 95% CI:

[-.321, -.094]; In Part II, based on 25 effect sizes, Fisher's z = -.150, *s.e.* = .051, p = .004, r = -.149, 95% CI_r: [-.246, -.049].

For Part I, Egger's regression test was non-significant at $p > .10$, $t(7) = -1.678$, $p = .137$. *Trim and Fill* procedure identified two missing studies on the right (the positive side). After filling in the two missing studies, the overall effect size based on 11 effect sizes was d = -.153, *s.e.* = .064, p = .017, 95% CI: [-.279, -.027]. The funnel plot with the two filled studies is displayed in Figure 2a. The adjusted effect size estimate was only slightly smaller, and still statistically significant. For Part II, Egger's test was non-significant at $p > .10$: $t(23) = 0.024$, p = .981. No missing studies were identified by *Trim and Fill*, therefore there was no adjustment of effect size estimate. The funnel plot is displayed in Figure 2b.[13] Using Borenstein et al.'s categories, we regard the potential impact of publication bias to be "trivial" for both analyses.

Discussion

Previous meta-analyses have shown that narratives can persuade (Braddock & Dillard, 2016; Shen et al., 2015; Tukachinsky & Tokunaga, 2013), but not whether this is due to reduced resistant message processing. One proposed theoretical mechanism is that narratives enhance persuasion by overcoming resistance (Dal Cin et al., 2004; Moyer-Gusé, 2008; Slater & Rouner, 2002). Empirical studies since examined this theoretical idea using experimental (comparing narratives vs. nonnarratives) and correlational (between narrative engagement and resistance) data. In this paper, we reported a two-part meta-analysis that synthesized these extant findings, with the goal of answering two questions: (1) Do narratives overcome resistance, and if so, what is the effect size? and (2) What factors, if any, may affect the magnitude of the effect?

What is the Overall Effect?

In answer to the first question, according to both analyses, narratives did help overcome resistance. There was less resistance, either as a result of viewing narrative vs. nonnarrative messages (Part I: d = -.214, equivalent r = -.106, $p < .001$), or as a correlate with greater engagement with narrative messages (Part II: r = -.131, equivalent d = -.264, $p < .001$). The effects were both in the "small" category (Cohen, 1992). Findings from experimental data and correlational data corroborate theoretical arguments about resistance attenuation in the context of narrative persuasion (Green & Brock, 2002; Moyer-Gusé, 2008; Moyer-Gusé & Nabi, 2010; Slater & Rouner, 2002). Though the relationship between resistance and persuasive outcomes was not part of our analyses (it has been demonstrated in previous meta-analytical studies, such as Rains, 2013), our findings add much empirical credence to resistance reduction as a causal pathway between exposure to narratives and persuasive outcomes.

Does the Effect Vary, and How?

Our analyses showed evidence of heterogeneity in both Part I and II, suggesting the presence of moderator variables. Part II data allowed us to examine a host of moderators that may potentially explain the heterogeneity (see Table 5). Among the sample and study characteristics, we found that the sample average age, female percentage, and data collection setting emerged as significant moderators. Especially worth noting is the rather large effect of gender composition: A larger effect size was observed for samples with a higher percentage of female participants. Given that primary studies in this area of research had an overall majority of female participants (the average percentage of female participants was 65.41% and 68.95% for Part I & II respectively), there could be a female bias in the observed overall effect sizes. In future research, researchers should consider achieving better gender balance in their sampling approach.

Narrative message characteristics. Most notably, our findings point to the importance of narrative message characteristics. In Part II data, including message characteristics in the model reduced the between-study heterogeneity to zero. Specifically, we observed a larger effect size for (education-) entertainment compared to advocacy messages. This finding was in alignment with the argument that narrative effects may be contingent upon the explicitness of the persuasive effort (Bilandzic & Busselle, 2013). We also found that messages that were of medium length (e.g., a 15- to 20-minute TV program) generated a larger effect size compared to both short (e.g., a 30-second PSA or a one-page story) and long (e.g., a 30- or 40-minute TV program) messages. Short messages may not be sufficient to get the participants engaged with the narrative, while long ones could produce fatigue or introduce other nuisance factors that may weaken the relationship. The effect size was also larger when the program featured one primary character to identify with rather than multiple characters. Featuring a single character might facilitate more concentrated narrative processing, with greater immersion into the storyline and/or identification with the character; it could also potentially reduce measurement noise by avoiding participants' confusion over with "whom" to identify or counterargue when they answer the related survey questions. These findings have practical implications for message design. For example, everything else equal, narrative messages that are of medium length, that better disguise persuasive intent, or that highlight one identifiable character should be less likely to arouse resistance.

Medium also had a significant effect: Textual materials showed a stronger negative relationship between narrative engagement and resistance compared to audio and video formats. This seems somewhat contrary to Shen et al.'s (2015) meta-analytic finding, which showed that narrative effects on persuasive outcomes were significant for audio and video messages, but not

messages in print. One possible explanation is that Shen et al.'s findings were about the role of medium in amplifying or weakening the *relative* effect of narratives over nonnarratives. In our Part II data, the coefficient of medium represents the comparison between audiovisual vs. textual format *within* narrative messages after controlling for other message characteristics. In other words, for example, greater engagement with a one-page textual message featuring one character, compared to a 50-second PSA with one main character, would be related to less resistance. As the short audiovisual narratives used in the studies in Part II were mostly PSAs, the audiovisual impact could render the message more engaging but at the same time more "in-the-face," thus generating greater resistance.

From a theoretical standpoint, these findings show that not all narratives are created equal, highlighting the need for more systematic explication and testing of narrative message features. The current landscape of narrative research has amassed a wide range of message stimuli lumped under the umbrella term "narratives." Not much attention has been paid to *a priori* explication of message features and theory-based designs to specifically investigate *what* message dimensions may enhance engagement or suppress resistance, and under what boundary conditions. Future studies should undertake conceptual explications of narrative message features, and experimentally manipulate levels of narrative engagement to shed better light on the psychological mechanisms of narrative persuasion. We see this as an important next step in narrative research for both theoretical advancement and practical implications for effective message design.

Construct operationalization. As noted earlier, the idea of "narratives overcoming resistance" was examined in extant research with a variety of constructs of resistance and narrative engagement. In response to RQ1, findings from Part II showed significant differences

between anger and other forms of resistance, such that the relationship between engagement and resistance was the weakest for anger. The average effect sizes for the subcategories (Table 4) showed that the relationship was significant for all cognitive forms of resistance, but not for anger. There was also a significant interaction effect: With freedom threat, transportation had a stronger negative relationship than character-based engagement (identification and parasocial interaction), whereas the reverse was true for reducing anger. Being transported into the storyline, in other words, did not seem to be effective in reducing the emotional form of resistance, whereas identifying or (parasocially) interacting with a character helped overcome anger but did not decrease the perception of threatened freedom. As a group, these construct operationalization variables reduced the effect-size level variance by a small (though not statistically significant) amount.

In our analyses, these forms of resistance and processes of narrative engagement were treated as different operationalizations of the underlying constructs central to EORM (Moyer-Gusé, 2008) as this approach best represents the primary studies. For example, some studies used a single scale focused on perceived freedom threat to indicate reactance (e.g., Quintero Johnson & Sangalang, 2017; Reinhart & Anker, 2012), or measured only negative cognitive responses (e.g., Keer et al., 2013). Whereas the primary goal of our syntheses is to test EORM's central proposition about resistance, we acknowledge the interrelationships among freedom threat, anger, and negative cognition as established in the reactance literature (Dillard & Shen, 2005; Rains, 2013), which we are unable to empirically specify and test as such in our data due to the lack of primary data. For the same reason, we could not address the interrelationships among narrative engagement types in our analyses, the incorporation of which would yield a more precise estimate of the unique contributions of each engagement type.

A Note on Power

The power of detecting moderator effects in a meta-analysis is often very low (Hedges & Pigott, 2004). While highlighting the above significant findings, we also hasten to emphasize that the absence of statistical significance does not necessarily mean no effect. Assessing power for a meta-analysis is a challenging task, especially for random-effects models, as prospective power analysis (which is recommended over retrospective power analysis) requires assumptions about parameters that are unknown before the review (especially for a random-effects model; Borenstein et al, 2009; Valentine et al., 2010). Here we echo Valentine et al.'s (2010) suggestion that meta-analytic findings are best interpreted using both the point estimate and the confidence interval, which provides information about the range of values and the amount of uncertainty in the estimate. Confidence intervals should be used in addition to or in place of the *p* value and can be more informative than the power analysis alone (Valentine et al., 2010). We reported confidence intervals for all the analyses in our paper, and hope that readers use that information in interpreting the results.

Limitations

Our study has a few limitations. First, the number of studies included in Part I was small. We were only able to locate 15 effect sizes that directly compared narrative and nonnarrative messages on some measures of message resistance. Whereas a meta-analytic review of this size is not uncommon (Borenstein et al., 2009)[14] and it is adequate for achieving reasonable power in random-effects models (Jackson & Turner, 2017), the small number of effect sizes prevented us from a moderator analysis, and the publication bias analysis could also be under-powered. Though Egger's regression test was non-significant, the *Trim and Fill* procedure detected two missing studies (see the funnel plot, Figure 2a), yielding a slightly smaller but still significant

adjusted mean effect size. Experimental designs will allow more direct testing of the theorized mechanisms and increase confidence in causal inferences about the effect of narratives. The small corpus of extant experiment studies calls for more future research that uses experimental methods to investigate mechanisms underlying narrative persuasion.

Second, due to the lack of primary data on the interrelations among constructs, we could not investigate narrative engagement as a mediating path between exposure to a narrative and resistance. As a result, our two-part analysis contributed two separate pieces of evidence of narrative overcoming resistance—1) the direct experimental effect and 2) the correlation between narrative engagement and resistance—but not evidence about narrative message reducing resistance *via* narrative engagement processes.

Finally, we hope to draw narrative researchers' attention to exploring other potential theoretical conditions for narrative persuasion. For example, one such condition postulated by Slater and colleagues is the relationship between one's pre-existing attitudes and the position advocated by the message. In their explication of E-ELM, they wrote, "absorption in a narrative, and response to characters in a narrative, should enhance persuasive effects and suppress counterarguing *if the implicit persuasive content is counterattitudinal*" (Slater & Rouner, 2002, p. 173, italics added). Echoing that, Dal Cin et al. (2004) argue that the persuasive advantage of narratives should be specifically in changing "strong attitudes—those that truly elicit resistance" (p. 177). From our review of the literature, this point has not received its due attention from narrative persuasion researchers. More primary data on such theoretically-based conditions would allow future syntheses to better delineate and evaluate theoretical developments in this area of research.

Conclusion

This paper reports a two-part meta-analysis that synthesized experimental and correlational data on narratives overcoming resistance. Results from Part I suggested that narratives, compared to nonnarratives, generated less resistance. Part II showed narrative engagement was negatively correlated with resistance, and narrative message characteristics were found to be important moderators of the relationship. Results from these syntheses help to resolve what appeared as inconsistent or contradictory findings in some primary studies. As was predicted in EORM, narrative messages can indeed be an effective strategy in reducing audience resistance. The summary effect sizes generated in this meta-analysis will also serve as useful yardsticks for the calculation of sample size and power in future studies.

References

Note: References marked with an asterisk (*) are included in the meta-analysis.

Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, *14*, 125–131. doi:10.1080/08824099709388654

Appel, M., & Malečkar, B. (2012). The influence of paratext on narrative persuasion: Fact, fiction, or fake? *Human Communication Research*, *38*, 459–484. doi:10.1111/j.1468-2958.2012.01432.x

*Asbeek Brusse, E. D., Neijens, P. C., & Smit, E. G. (2010, December). Talking to deaf ears? Two routes to narrative persuasion in entertainment education. In *International Conference on Future Imperatives of Communication and Information for Development and Social Change, Bangkok.*

*Banerjee, S. C., & Greene, K. (2012). Role of transportation in the persuasion process: Cognitive and affective responses to antidrug narratives. *Journal of Health Communication*, *17*, 564–581. doi:10.1080/08870446.2012.675063

Becker, B. J. (2006). Failsafe N or file-drawer number. In Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 111–126). John Wiley & Sons.

Bilandzic, H., & Busselle, R. W. (2013). Narrative persuasion. In J. P. Dillard & L. J. Shen (Eds.), *The Sage handbook of persuasion: Developments in theory and practice* (pp. 200–219). Los Angeles, London: Sage. doi:10.4135/9781452218410.n13

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley. doi:10.1002/9780470743386

Braddock, K., & Dillard, J. P. (2016). Meta-analytic evidence for the persuasive effect of narratives on beliefs, attitudes, intentions, and behaviors. *Communication Monographs*, *83*, 446–467. doi:10.1080/03637751.2015.1128555

Brehm, S. S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control*. New York, NY: Academic Press.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*, 1–21. doi:10.1086/448619

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*, 321–347. doi:10.1080/15213260903287259

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*, 1531–1546. doi:10.1177/0956797617714579

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, *4*, 245–264. doi:10.1207/S15327825MCS0403_01

Cohen, J. (2006). Audience identification with media characters. In J. Bryant and P. Vorderer (Eds.), *Psychology of entertainment* (pp. 183–194). Mahwah, NJ: Lawrence Erlbaum.

Cohen, J., & Tal-Or, N. (2017). Antecedents of identification. In F. Hakemulder, M. M. Kuijpers, E. S. Tan, K. Bálint, and M. M. Doicaru (Eds.), *Narrative absorption*. Amsterdam, The Netherlands: John Benjamins Publishing Company. doi:10.1075/lal.27.08coh

Dal Cin, S., Zanna, M.P., Fong, G.T., 2004. Narrative persuasion and overcoming resistance. In: Knowles, E.S., Linn, J.A. (Eds.), *Resistance and persuasion* (pp. 175–191). Lawrence Erlbaum, Mahwah, NJ.

*de Graaf, A., van den Putte, B., Nguyen, M. H., Zebregs, S., Lammers, J., & Neijens, P. (2017). The effectiveness of narrative versus informational smoking education on smoking beliefs, attitudes and intentions of low-educated adolescents. *Psychology & Health*, *32*, 810–825. doi:10.1080/08870446.2017.1307371

Dillard, J. P., & Peck, E. (2001). Persuasion and the structure of affect: Dual systems and discrete emotions as complementary models. *Human Communication Research*, *27*, 38–68. doi:10.1111/j.1468-2958.2001.tb00775.x

Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, *72*, 144–168. doi:10.1080/03637750500111815

Dopp, A. R., Borduin, C. M., White, I. I., Mark, H., & Kuppens, S. (2017). Family-based treatments for serious juvenile offenders: A multilevel meta-analysis. *Journal of Consulting and Clinical Psychology*, *85*, 335–354. doi:10.1037/ccp0000183

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–99. doi:10.1080/01621459.2000.10473905

Escalas, J. E. (2004). Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology, 14,* 168–180.

*Gardner, L., & Leshner, G. (2016). The role of narrative and other-referencing in attenuating psychological reactance to diabetes self-care messages. *Health Communication*, *31*, 738–751. doi:10.1080/10410236.2014.993498

Givens, G. H., Smith, D. D., & Tweedie, R. L. (1997). Publication bias in meta-analysis: a Bayesian data-augmentation approach to account for issues exemplified in the passive smoking debate. *Statistical Science*, *12*, 221–240.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*, 701–721. doi:10.1037/0022-3514.79.5.701

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, and T. C. Brock (Eds.) *Narrative impact: Social and cognitive foundations* (pp. 315–41). Mahwah, NJ: Lawrence Erlbaum Associates.

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, *14*, 311-327. doi:10.1111/j.1468-2885.2004.tb00317.x

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. doi:10.1080/19312450709336664

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, *9*, 426–445.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*, 486–504. doi:10.1037/1082-989X.3.4.486

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0). Retrieved from http://www.cochrane.org/handbook

Hinyard, L. J., & Kreuter, M. W. (2007). Using narrative communication as a tool for health behavior change: A conceptual, theoretical, and empirical overview. *Health Education & Behavior*, *34*, 777–792. doi:10.1177/1090198106291963

Hox, J. (2002). The multilevel approach to Meta-analysis. In J. Hox. *Multilevel analysis: Techniques and applications*. Mahwah, NJ: LEA.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative knowledge in psychology. *International Journal of Selection and Assessment*, *8*, 275–292. doi:10.1111/1468-2389.00156

*Igartua, J. J., & Vega Casanova, J. (2016). Identification with characters, elaboration, and counterarguing in entertainment-education interventions through audiovisual fiction. *Journal of Health Communication*, *21*, 293–300. doi:10.1080/10810730.2015.1064494

*Keer, M., van den Putte, B., de Wit, J., & Neijens, P. (2013). The effects of integrating instrumental and affective arguments in rhetorical and testimonial health messages. *Journal of Health Communication*, *18*, 1148–1161. doi:10.1080/10810730.2013.768730

*Kim, H. K., & Niederdeppe, J. (2016). Effects of self-affirmation, narratives, and informational messages in reducing unrealistic optimism about alcohol-related problems among college students. *Human Communication Research*, *42*, 246–268. doi:10.1111/hcre.12073

Knowles, E. S., & Linn, J. A. (Eds.). (2004). *Resistance and persuasion*. Mahwah, NJ: Psychology Press.

*Krakow, M., Jensen, J. D., & Christy, K. (2017, May). *Life or death in cervical cancer narratives: Testing the mediating effects of transportation, reactance, and message impact.* Paper presented at the Preconference on Innovations in Narrative-Based Intervention

Design, Development, and Evaluation at the annual meeting of the International Communication Association, San Diego, CA.

*Krakow, M., Yale, R. N., Jensen, J. D., Carcioppolo, N., & Ratcliff, C. L. (2018). Comparing mediational pathways for narrative- and argument-based messages: Believability, counterarguing, and emotional reaction. *Human Communication Research*, *44*, 299–321. doi:10.1093/hcr/hqy002.

*Krakow, M., Yale, R. N., Perez Torres, D., Christy, K., & Jensen, J. D. (2017). Death narratives and cervical cancer: Impact of character death on narrative processing and HPV vaccination. *Health Psychology*, *36*, 1173–1180. doi:10.1037/hea0000498

*Kreuter, M. W., Holmes, K., Alcaraz, K., Kalesan, B., Rath, S., Richert, M., ... & Clark, E. M. (2010). Comparing narrative and informational videos to increase mammography in low-income African American women. *Patient Education and Counseling*, *81*, S6–S14. doi:10.1016/j.pec.2010.09.008

Lindsey, L. L. M. (2005). Anticipated guilt as behavioral motivation an examination of appeals to help unknown others through bone marrow donation. *Human Communication Research*, *31*, 453-481. doi:10.1111/j.1468-2958.2005.tb00879.x

Lombard, M., Reich, R. D., Grabe, M. E., Bracken, C. C., & Ditton, T. B. (2000). Presence and television. *Human Communication Research*, *26*(1), 75–98. doi:10.1111/j.1468-2958.2000.tb00750.x

Madden, T. J., Allen, C. T., & Twible, J. L. (1988). Attitude toward the ad: An assessment of diverse measurement indices under different processing “sets”. *Journal of Marketing Research*, *25*, 242–252. doi:10.2307/3172527

*McQueen, A., Kreuter, M. W., Kalesan, B., & Alcaraz, K. I. (2011). Understanding narrative effects: The impact of breast cancer survivor stories on message processing, attitudes, and

beliefs among African American women. *Health Psychology*, *30*, 674–682. doi:10.1037/a0025395

Moyer-Gusé, E. (2008). Toward a theory of entertainment persuasion: Explaining the persuasive effects of entertainment-education messages. *Communication Theory*, *18*, 407–425. doi:10.1111/j.1468-2885.2008.00328.x

*Moyer-Gusé, E., Chung, A. H., & Jain, P. (2011). Identification with characters and discussion of taboo topics after exposure to an entertainment narrative about sexual health. *Journal of Communication*, *61*, 387–406. doi:10.1111/j.1460-2466.2011.01551.x

Moyer-Gusé, E., Jain, P., & Chung, A. H. (2012). Reinforcement or reactance? Examining the effect of an explicit persuasive appeal following an entertainment-education narrative. *Journal of Communication*, *62*, 1010–1027.

*Moyer-Gusé, E., & Nabi, R. L. (2010). Explaining the effects of narrative in an entertainment television program: Overcoming resistance to persuasion. *Human Communication Research*, *36*, 26–52. doi:10.1111/j.1468-2958.2009.01367.x

Nabi, R. L., Moyer-Gusé, E., & Byrne, S. (2007). All joking aside: A serious investigation into the persuasive effect of funny social issue messages. *Communication Monographs*, *74*(1), 29-54. doi:10.1080/03637750701196896

*Niederdeppe, J., Heley, K., & Barry, C. L. (2015). Inoculation and narrative strategies in competitive framing of three health policy issues. *Journal of Communication*, *65*, 838–862. doi:10.1111/jcom.12162

*Niederdeppe, J., Kim, H. K., Lundell, H., Fazili, F., & Frazier, B. (2012). Beyond counterarguing: Simple elaboration, complex integration, and counterelaboration in response to variations in narrative focus and sidedness. *Journal of Communication*, *62*, 758–777. doi:10.1111/j.1460-2466.2012.01671.x

Niederdeppe, J., Roh, S., & Dreisbach, C. (2016). How narrative focus and a statistical map shape health policy support among state legislators. *Health Communication*, *31*, 242-255.

*Niederdeppe, J., Roh, S., & Shapiro, M. A. (2015). Acknowledging individual responsibility while emphasizing social determinants in narratives to promote obesity-reducing public policy: A randomized experiment. *PLoS ONE*, *10*, e0117565. doi:10.1371/journal.pone.0117565

*Niederdeppe, J., Shapiro, M. A., Kim, H. K., Bartolo, D., & Porticella, N. (2014). Narrative persuasion, causality, complex integration, and support for obesity policy. *Health Communication*, *29*, 431–444. doi:10.1080/10410236.2012.761805

*Niederdeppe, J., Shapiro, M. A., & Porticella, N. (2011). Attributions of responsibility for obesity: Narrative communication reduces reactive counterarguing among liberals. *Human Communication Research*, *37*, 295–323. doi:10.1111/j.1468-2958.2011.01409.x

*Quintero Johnson, J. M., & Sangalang, A. (2017). Testing the explanatory power of two measures of narrative involvement: An investigation of the influence of transportation and narrative engagement on the process of narrative persuasion. *Media Psychology*, *20*, 144–173. doi:10.1080/15213269.2016.1160788

Pastor, D. A., & Lazowski, R. A. (2018). On the multilevel nature of meta-analysis: A tutorial, comparison of software programs, and discussion of analytic choices. *Multivariate Behavioral Research*, *53*(1), 74-89. doi:10.1080/00273171.2017.1365684

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rains, S. A. (2013). The nature of psychological reactance revisited: A meta-analytic review. *Human Communication Research*, *39*(1), 47-73. doi:10.1111/j.1468-2958.2012.01443.x

Raudenbush, S.W., & Bryk, A.S., (2002). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

*Reinhart, A. M., & Anker, A. E. (2012). An exploration of transportation and psychological reactance in organ donation PSAs. *Communication Research Reports*, *29*, 274–284. doi:10.1080/08824096.2012.704601

Rubin, A. M., & Perse, E. M. (1987). Audience activity and soap opera involvement: A uses and effects investigation. *Human Communication Research*, *14*, 246–268. doi:10.1111/j.1468-2958.1987.tb00129.x

*Sangalang, A. L. (2015). Using stories to communicate electronic cigarette risk to young adults: Explicating similarity in narrative persuasion (Dissertation). Retrieved from ProQuest Information & Learning. (10014569)

Scargle, J. D. (1999). Publication bias (the "file-drawer problem") in scientific inference. *arXiv preprint physics/9909033*. doi:10.1037/1082-989X.10.4.428

*Scherr, C. L., Jensen, J. D., Krakow, M., Ratcliff, C. L., & Liu, M. (2017, May). Modeling psychological reactance using narratives from the AT&T "It Can Wait" campaign. Paper presented at the International Communication Association Annual Meeting, San Diego, CA.

Schmidt, F. L., Oh, I-S., & Hayes, T. L. (2009). Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results. *British Journal of Mathematical and Statistical Psychology*, *62*, 97–128. doi:10.1348/000711007X255327

Schramm, H., & Hartmann, T. (2008). The PSI-process scales: A new measure to assess the intensity and breadth of parasocial processes. *Communications-European Journal of Communication Research*, *33*, 385–401. doi:10.1515/COMM.2008.025

*Shade, Drew D. (2014). Television as an instrument of social change: Examining involvement, musical performance, and viewer effects (Dissertation). Retrieved from ProQuest Information & Learning. (10170387)

*Shen, L., Seung, S., Andersen, K. K., McNeal, D. (2017). The psychological mechanisms of persuasive impact from narrative communication. *Studies in Communication Sciences*, *2*, 165–181. doi:10.24434/j.scoms.2017.02.003

Shen, F., Sheer, V. C., & Li, R. (2015). Impact of narratives on persuasion in health communication: A meta-analysis. *Journal of Advertising*, *44*, 105–113. doi:10.1080/00913367.2015.1018467

Slater, M. D., & Rouner, D. (2002). Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory*, *12*, 173–191.

Sterne, A. C. & Egger, M. (2006). Regression methods to detect publication and other bias in meta-analysis. In Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). *Publication bias in meta-analysis: Prevention, assessment and adjustments, pp. 99-110*. John Wiley & Sons.

Tal-Or, N., & Cohen, J. (2010). Understanding audience involvement: Conceptualizing and manipulating identification and transportation. *Poetics*, *38*, 402–418.

Tukachinsky, R., & Tokunaga, R. S. (2013). The effects of engagement with entertainment. *Communication Yearbook*, 37, 287–321. doi:10.1080/23808985.2013.11679153

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*, 215–247. doi:10.3102/1076998609346961

van Assen, M. A., van Aert, R., & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, 20, 293–309.

van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, *40*, 797–817. doi:10.1086/673383

Vevea, J. L., & Woods, C. M. (2005). Publication Bias in Research Synthesis: Sensitivity Analysis Using A Priori Weight Functions. *Psychological Methods, 10*, 428-443.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. http://www.jstatsoft.org/v36/i03/

Vorderer, P. A., & Hartmann, T. (2009). Entertainment and enjoyment as media effects. In Bryant, J., & Oliver, M. B. (Eds.). *Media effects: Advances in theory and research, 532-550*. Routledge.

*Walter, N., Murphy, S. T., Frank, L. B., & Baezconde-Garbanati, L. (2017). Each medium tells a different story: The effect of message channel on narrative persuasion. *Communication Research Reports*, *34*, 161–170. doi:10.1080/08824096.2017.1286471

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communications Monographs*, *59*, 329–349. doi:10.1080/03637759209376276

Witte, K. (1994). Fear control and danger control: A test of the extended parallel process model (EPPM). *Communications Monographs*, *61*, 113–134. doi:10.1080/03637759409376328

Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: A meta-analysis. *Health Communication*, *30*, 282–289. doi:10.1080/10410236.2013.842528

*Zhou, S., & Shapiro, M. A. (2017). Reducing resistance to narrative persuasion about binge drinking: The role of self-activation and habitual drinking behavior. *Health Communication*, *32*, 1297–1309. doi:10.1080/10410236.2016.1219931

Table 1a.

Summary of Resistance Types in Included Papers (Part I)

	Counterarguing	Anger	Freedom Threat	Message Derogation
de Graaf et al. (2017)*	x			
Gardner & Leshner (2016)*	x	x	x	
Keer et al. (2013)*				x
Kim & Niederdeppe (2016)*				x
Krakow et al. (2018), studies 1 and 2*	x	x	x	
Kreuter et al. (2010)	x			
Niederdeppe et al. (2011)*	x			
Niederdeppe, Heley et al. (2015)*	x			

* Additional data were requested and obtained from the authors to compute our effect sizes.

Table 1b.

Summary of Resistance & Narrative Engagement Types in Included Papers (Part II)

	Counterarguing	Anger	Freedom Threat	Message Derogation	Transportation	Identification	Parasocial Interaction
Asbeek Brusse et al. (2010)	x				x	x	
Banerjee & Greene (2012)		x			x		
Igartua & Vega Casanova (2016)*	x					x	
Keer et al. (2013)				x	x		
Kim & Niederdeppe (2016)*				x	x	x	
Krakow, Jensen et al. (2017)*	x	x	x		x		
Krakow, Yale et al. (2017)*	x	x	x		x	x	
McQueen & Kreuter (2010)				x	x		
Moyer-Gusé & Nabi (2010)	x		x		x	x	x
Moyer-Gusé et al. (2011)	x					x	
Niederdeppe, Roh et al. (2015)	x					x	
Niederdeppe et al. (2012)*	x					x	
Niederdeppe et al. (2014)*	x	x				x	
Quintero Johnson & Sangalang (2017)	x		x		x		
Reinhart & Anker (2012)			x		x		
Sangalang (2015)*	x		x		x		
Scherr et al. (2017)*		x	x	x	x		
Shade (2014)	x				x	x	
Shen et al. (2017)*		x	x	x	x	x	x
Walter et al. (2017), studies 1 and 2*			x		x		
Zhou & Shapiro (2017), studies 1 and 2*	x		x	x	x	x	

* Additional data were requested and obtained from the authors to compute our effect sizes.

Table 2.

Moderator Variables and the Number of Effect Sizes Associated with Each Variable/Category

Variables	Categories and Definitions	Part I (*k*)	Part II (*k*)
Sample demographic characteristics			
Mean age	Average age of the participants in a sample	14	50
Female %	Percentage of female participants in a sample	15	63
White %	Percentage of White Participants in a sample	13	50
Study/report characteristics			
Data Collection Setting	*Lab* (researchers' computer lab)	4	28
	Online distribution and collection (e.g., M-Turk)	9	31
	Other (e.g., a shopping mall)	2	4
Sample	*College Students*	2	37
	General Population	8	9
	Specific Segment (e.g., adolescents, African-American women, patients with diabetes)	5	17
Publication Status	*Unpublished*	0	12
	Published	15	51
Narrative Stimuli Characteristics			
Length	*Short* (e.g., 400-word article, 30-second PSA, 1-page comic strip)	14	35
	Medium (e.g., 20-minute TV program)	0	11
	Long (e.g., 45-minute program)	1	17
Genre	*Advocacy* (content containing explicit advocacy for a specific change, e.g., PSAs)	15	42
	(Education-) Entertainment (content created to entertain and/or educate audience, e.g., TV shows such as *Sex and the City*, or *Grey's Anatomy*)	0	21
Characters	*Single* (story contains a single main character to identify with; e.g., Zhou & Shapiro, 2017)	7	32
	Multiple (story contains multiple main characters; e.g., Quintero Johnson & Sangalang, 2017)	8	31
Medium	*Audiovisual* (content is watched and/or listened to, e.g., a video)	7	43
	Textual (content is read, e.g., a booklet)	7	20
	Mixed (e.g., text plus audio recording; Niederdeppe et al., 2011)	1	0
Construct Operationalization			
Narrative Engagement	*Transportation*	N/A	33
	Character Identification	N/A	25
	Parasocial Interaction	N/A	5
Resistance	*Anger*	3	9
	Freedom Threat	3	19
	Counterarguing	7	23
	Message Derogation	2	12

Table 3.

Mean Effect Sizes and Heterogeneity from Models Without Moderators

			Summary Effect Size			Heterogeneity Indices		
	k	Total N	Mean Effect Size[b]		Confidence Interval	*Q*(*df*)	I^2	τ^2
Part I			*d (s.e.)*		95% CI_d			
FE model			-.168 (.020)		[-.207, -.128]		--	--
2-level RE model	15	5,397	-.200 (.035)		[-.279, -.130]	28.69* (14)	43.29%	.006
3-level RE model[a]			-.213 (.054)		[-.318, -.107]		66.77%	.015
Part II			Fisher's *z*	*r*	95% CI_r			
FE model			-.123 (.008)	-.122	[-.137, -.107]		--	--
2-level RE model	63	17,184	-.113 (.025)	-.113	[-.161, -.064]	659.80*** (62)	90.51%	.035
3-level RE model[a]			-.131 (.038)	-.131	[-.203, -.057]		91.61%	.040

Notes.
*$p < .05$, *** $p < .001$
FE = Fixed-effect, RE = Random-effects, *k* = the number of effect sizes for each analysis
[a] The number of clusters (i.e., studies) at Level 3 was 9 for Part I, and 25 for Part II.
[b] All significant at $p < .001$.

Table 4.

Mean Effect Size by Resistance Type and Narrative Engagement Type (Part II)

	k	Fisher *z*	*(s.e.)*	95% CI_z	*r*
By Resistance Type					
Anger	9	-.063	(.064)	[-.189, .062]	-.063
Freedom threat	19	-.110*	(.050)	[-.208, -.012]	-.110*
Counterarguing	23	-.135**	(.047)	[-.228, -.042]	-.134**
Message Derogation	12	-.213***	(.061)	[-.332, -.094]	-.210***
By Narrative Engagement Type					
Transportation	33	-.149***	(.044)	[-.236, -.063]	-.148***
Identification	25	-.107*	(.048)	[-.201, -.013]	-.107*
Parasocial relationship	5	-.137	(.088)	[-.310, .036]	-.136

Notes.
*$p < .05$, **$p < .01$, ***$p < .001$
Estimates were obtained from three-level random-effects models.
k = the number of effect sizes for each analysis

Table 5.

Results of Three-level Mixed-effects Models for Moderator Analyses (Part II)

		b (s.e.)	95% CI	Model-Level Statistics
Models 1a- 1c: Sample Demographics				
Mean age (years)		-.008* (.004)	[-.015, -.001]	
Female %		-.524** (.164)	[-.845, -.202]	
White %		.014 (.154)	[-.288, .316]	
Model 2: Study/Report Characteristics				
Setting *(Lab)*	*Online*	-.057 (.083)	[-.219, .106]	
	Natural	-.294* (.144)	[-.576, -.013]	
Sample *(College)*	*General*	.136 (.127)	[-.113, .385]	
	Specific	.136 (.086)	[-.033, .304]	Q_M (*df*): 5.17(5)
Publication *(Unpublished)*	*Published*	.040 (.090)	[-.137, .216]	$R^2_{META(2)}$: 2.87%
				$R^2_{META(3)}$: 22.31%
Model 3: Narrative Message Characteristics				
Length *(Short)*	*Medium*	-.408*** (.114)	[-.631, -.185]	
	Long	-.124 (.143)	[-.405, .157]	
Genre (*Advocacy)*	*E-E*	-.350*** (.102)	[-.550, -.149]	
Characters *(Single)*	*Multiple*	.406*** (.091)	[.227, .584]	
				Q_M (*df*): 42.01(5)***
Medium [a] *(Audiovisual)*	*Textual*	-.222*** (.052)	[-.323, -.120]	$R^2_{META(2)}$: 2.87%
				$R^2_{META(3)}$: 22.31%
Model 4: Construct Operationalization				
NE *(Content-based)*	*Character -based*	-.148 (.096)		
Resistance *(Anger)*	*FT*	-.163* (.078)	[-.316, -.010]	
	CA	-.164* (.083)	[-.327, -.001]	
	MD	-.241** (.083)	[-.416, -.065]	
Interaction terms	*NE* FT*	.264* (.115)	[.038, .489]	Q_M (*df*): 11.49(7)
	NE CA*	.183 (.113)	[-.039, .404]	$R^2_{META(2)}$: 27%
	NE MD*	.197 (.125)	[-.048, .442]	$R^2_{META(3)}$: 0%

Notes.
$^{\#}p < .10$, $*p < .05$, $** p < .001$, $*** p < .001$

Unstandardized coefficients are reported in the table.

The parentheses under each variable name designates the reference category.

E-E= Education-Entertainment, NE= Narrative Engagement, FT= Freedom Threat, CA= Counterarguing, MD=Message Derogation

[a] The "mixed" category for this variable was omitted as it contained only one effect size for Part I, and did not apply to Part II.

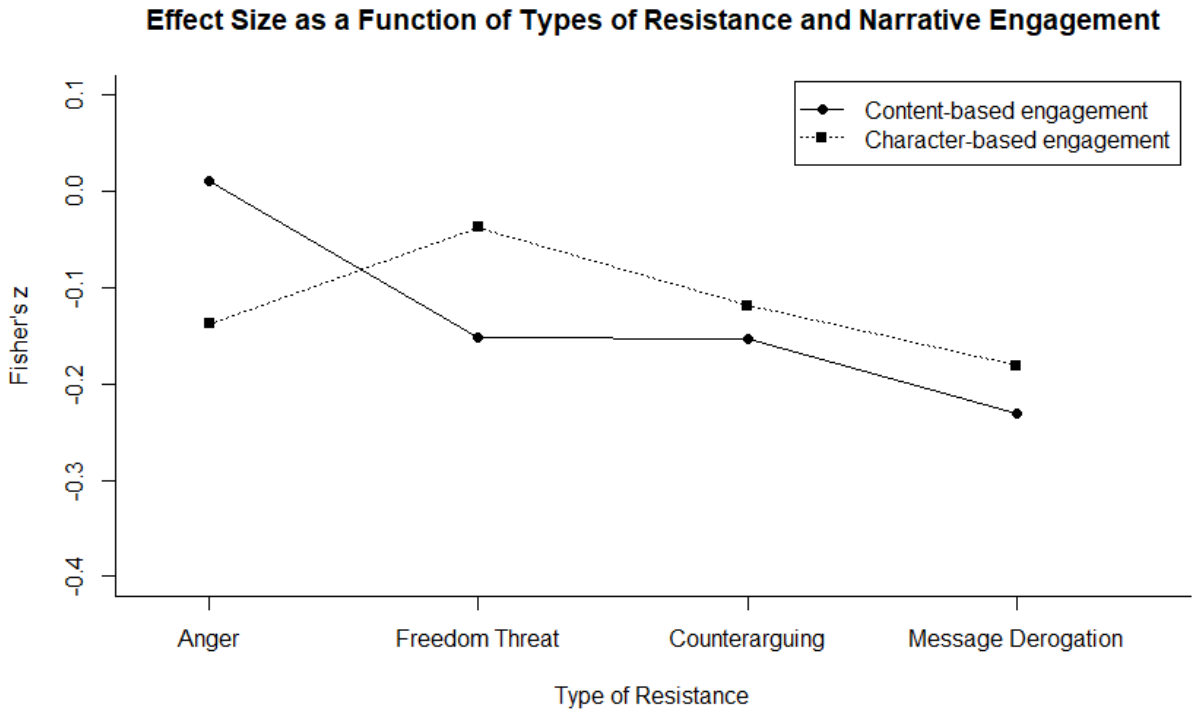

Figure 1. Depicting the Interaction Effect between Types of Reactance and Narrative Engagement.

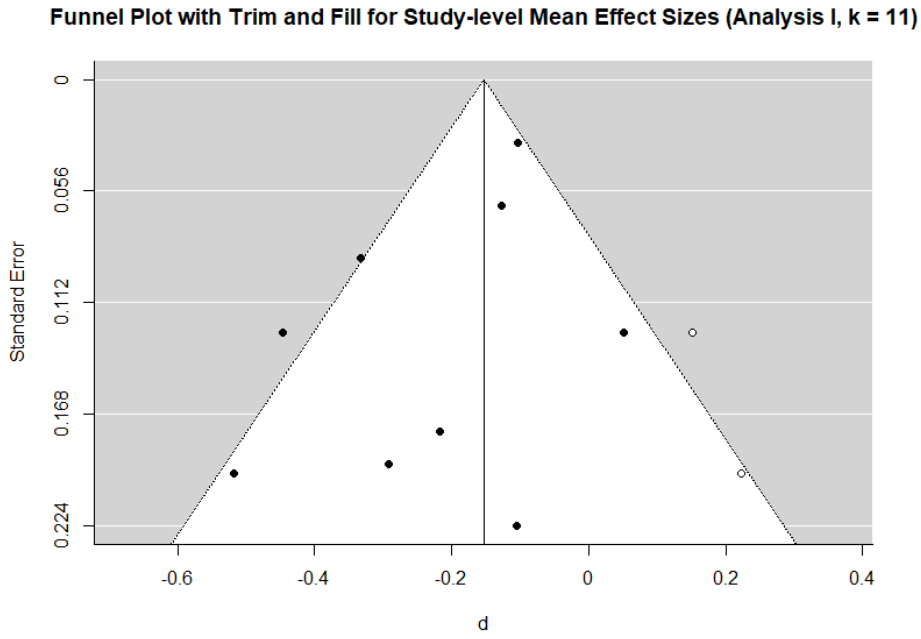

Figure 2a. Funnel Plot (with Trim and Fill) for Part I

Note. The two white circles on the right side represent the "filled" studies through the trim-and-fill method.

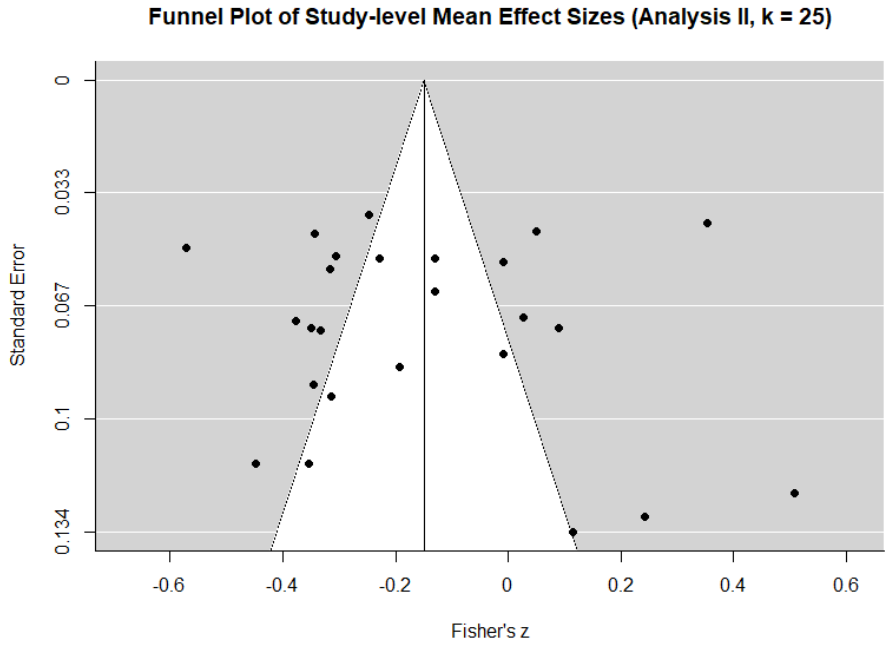

Figure 2b. Funnel Plot for Part II.

Figure 2. Funnel Plots of Study-Level Mean Effect Sizes

ENDNOTES

[1] Although questioning the veracity of a narrative, or "false note finding", is sometimes considered a form of critical thinking akin to counterarguing (e.g., van Laer et al., 2014; Green & Brock, 2000), we do not consider this to exclusively capture resistance. Questioning the plausibility of story events or character actions can certainly be a form of message discounting; however, we take into account that perceived fictionality does not necessarily mean rejection of a story or its underlying argument. Indeed, fictional and fake stories can still be persuasive (Appel & Malečkar, 2012). For this reason, we do not include false note finding as a resistance measure in our analyses.

[2] Tukachinsky and Tokunaga (2013) considered perceived similarity between oneself and story characters to be another measure of narrative engagement. However, perceived similarity is typically characterized as a cognitive assessment *about* a narrative whereas absorptive processes entail immersion *into* the narrative (Cohen, 2001; Moyer-Gusé, 2008). Perceived similarity has also been suggested as an antecedent to engagement (specifically, to identification; Cohen, 2001), but this has not been well-supported empirically (see Cohen & Tal-Or, 2017). Aligning with the notion of engagement as suspension of one's own perspective, we do not include perceived similarity as a construct of narrative engagement in our analysis.

[3] Some studies in our analysis used, as an additional measure, the narrative engagement scale developed by Busselle and Bilandzic (2009), a multidimensional scale that captures narrative understanding, attentional focus, narrative presence, and emotional engagement. In these cases, we chose to extract effect size estimates from the unidimensional scales of transportation to be conceptually consistent with other studies.

[4] We used the Boolean search term: (reactance OR counterpersua* OR counter-argu* OR "counter argu*" OR counterargu* OR "threat to freedom" OR "freedom threat") AND (narrative* OR stories OR story).

[5] Some papers were conference papers or manuscripts under review at the time of our search, but were later published and thus have a publication date after our cutoff.

[6] Studies were excluded from Part I if the narrative was compared to a control condition that was not a comparable nonnarrative. For example, some studies used a no-message-exposure control (e.g., Niederdeppe et al., 2016), while others used an alternate narrative on a different topic (e.g., Moyer-Gusé, Jain, & Chung, 2012). These were excluded as the experimental vs. control conditions no longer represented narrative vs. nonnarrative message features, which is the theoretical interest of this meta-analysis.

[7] We thank all the authors who provided data upon our requests. For a few papers, we also obtained additional data that were not reported in the published text. For example, in Shen et al. (2017), the analyses reported in their manuscript were based on a composite measure of "reactance" (i.e., combining anger and negative cognition). Given our specific interest, we requested and obtained separate zero-order correlations involving each of these two measures and included them as different effect estimates. As another note, in one instance, we obtained the needed correlation from the dataset that the authors made available online (i.e., Niederdeppe, Roh, & Shapiro, 2015).

[8] Only a couple of papers reported other forms of data (e.g., odds ratio, Niederdeppe et al., 2011, or correlation, Gardner & Leshner, 2016), and they were converted into *d* using the appropriate formula (Borenstein et al., 2009).

[9] The *rma.mv* command for *metafor* package in R, which we use to estimate the meta regression models, performs listwise deletion. Including all three demographic variables in one model would reduce the sample size from 15 to 11 in Part I, and from 63 to 37 in Part II.

[10] Preliminary analysis on the subgroup means of identification and parasocial interaction showed that they were not statistically different (Means for the identification and parasocial interaction subgroups were respectively -.107 and -.137; mean difference = -.030, *s.e.* = 086, $p = .726$).

[11] We did not report Rosenthal's fail-safe N given the various concerns with this approach, including its problematic assumptions, underestimation of the extent of bias especially in the presence of heterogeneity, and widespread misinterpretations of its results (Givens, Smith, & Tweedie, 1997; Scargle, 1999; Vevea & Woods, 2005). Meta-analysts have recommended against reporting Rosenthal's fail-safe N as a test of publication bias (Becker, 2006).

[12] If publication bias is present, this procedure iteratively trims the most extreme small studies and recomputes the effect size until the funnel plot reaches symmetry. To restore the variance, it then adds back the original studies that were removed as well as their "mirror images" (i.e., imputed effect sizes in the opposite direction).

[13] We also did the same analyses using the full sample for both analyses, which yielded similar results. For Part I, Egger's test was non-significant: $t(13) = -1.56$, $p = .142$. *Trim and Fill* identified and filled in two missing effect sizes, and the adjusted average effect size based on 17 effect sizes was $d = -.181$, *s.e.*, = .037, $p < .001$, 95% CI: [-.253, -.100]. For Part II, Egger's test was non-significant: $t(61) = -.668$, $p = .507$. *Trim and Fill* identified no missing studies to fill.

[14] In the Cochrane Database of Systematic Reviews, a database consisting of 3000 reviews, the median number of studies included in a review is six.