Watching Videos on a Smartphone: Do Small Screens Impair Narrative Transportation?

Markus Appel & Christoph Mengelkamp

University of Würzburg, Germany

Author addresses

Prof. Dr. Markus Appel and Dr. Christoph Mengelkamp, Human-Computer-Media Institute,

University of Würzburg, Oswald-Külpe-Weg 82, 97074 Würzburg, Germany.

markus.appel@uni-wuerzburg.de; christoph.mengelkamp@uni-wuerzburg.de

Acknowledgments

We thank Anna Anastasopoulos, Anna Thoma, Anne-Katrin Nagel, Bernadette Motschenbacher, Carolin Schwärzler, Helena Karl, Isabel Öchsler, Jana Graber, Jiyeon Yeo, Larissa Volk, Laura Erk, Lea-Teresa Luzner, Lisa Bäuerle, Lisa Zink, Lorena Scheurich, Nadine Ley, Nele Friedrich, Paula Cimander, Pauline Zahn, Sabrina Beck, Secil Kilicarslan, Teresa Schell, Theresa Politt, and Tina Schwabe for their support in conducting the experiments.

Accepted for publication in the journal *Media Psychology*

https://doi.org/10.1080/15213269.2021.2025109

Abstract

Smartphones are a preferred platform to access audiovisual stories. Prior theory and research suggest that using smaller screens could lead to a shallower narrative experience. In three experiments we examined the influence of screen size (smartphone vs. computer screen) on the experience of being transported into the world of the story (narrative transportation). We further examined interaction effects with manipulations meant to change transportation by means of reviews (Experiment 1, $N = 120$), consistency of main character information (Experiment 2, $N = 139$), and prior information meant to facilitate comprehension (Experiment 3, $N = 129$). Because our series of studies involved theoretically and practically relevant null hypotheses (i.e., screen size does not influence transportation), we added Bayes factor analyses to standard frequentist statistics. A mini meta-analysis was conducted to summarize the results. Taken together, the three experiments indicate that smaller screen size does not impair narrative transportation. Implications and future research are discussed.

Keywords: Smartphones; Narrative Transportation; Stories; Bayes; Screen Size

Watching Videos on a Smartphone: Do Small Screens Impair Narrative Transportation?

Stories have the power to influence how we think about the real world, other human beings, and ourselves (Gottschall, 2013; Pinker, 2007), particularly if we are deeply transported into the world of a narrative (Green et al., 2020; van Laer et al., 2014). Reading and watching stories can be a source of enjoyment and well-being (Dill-Shackleford et al., 2016), stories help humans develop social abilities (Mar, 2018), and stories could be a key to reducing inter-group bias (Wojcieszak et al., 2020) and to improving environmental behavior (Morris et al., 2019). A substantial part of the stories we encounter are presented in an audiovisual format (Ryan, 2012). Along with the popularity of social media, video streaming (e.g., Netflix), and video sharing services (e.g., Youtube), smartphones have become a preferred platform to access audiovisual content (Clark, 2018; van Kessel et al., 2019). Thus, a large percentage of stories watched worldwide are watched on a small screen (van Kessel et al., 2019). Smartphone screens themselves tend to get bigger, but they are still much smaller than TV and computer screens that had been the main means to watching videos for decades. Whereas the mode of access for audiovisual content has puzzled artists and theorists for a long time (e.g., Godard, 1997), the empirical evidence on the impact of small screens on viewers is limited (Cummings & Bailenson, 2016). If watching stories on a smartphone would impair viewers' experience and story processing – making the experience shallow and superficial – this could have massive downstream effects to individuals and societies.

Our focus here is on the impact of screen size on recipients' experience of being transported into the world of a story, a main mechanism underlying story impact (Gerrig, 1993; Green et al., 2020; Green & Brock, 2000; 2002). We conducted three experiments to study the influence of screen size (smartphone vs. computer screen) on narrative transportation and we examined interaction effects with manipulations meant to change transportation by means of reviews

(Experiment 1, Tiede & Appel, 2020), consistency of main character information (Experiment 2, Busselle & Bilandzic, 2008), and prior information meant to facilitate understanding (Experiment 3, Green et al., 2008). If we found such interaction effects, researchers in basic and applied contexts would need to take into account screen size whenever manipulations of transportation were planned. Given the limitations of null hypothesis significance testing whenever the null hypothesis is a theoretically feasible possibility (e.g., Jarosz & Wiley, 2014; Wagenmakers, 2007), our results were analyzed with the help of Bayes factor analyses. Results were integrated with the help of a mini meta-analysis.

Experiencing Narrative Worlds

Much of the mediated information we encounter on a daily basis is presented in the form of a story or narrative. When following stories, individuals are often absorbed in the story world, they focus their attention on the story (rather than the surroundings in the place they are physically located) and have strong emotions (Gottschall, 2013; Oatley, 2016). In recent years, several constructs or concepts have been used to describe the state of being psychologically absorbed or immersed in a story (Hamby et al., 2018). Whereas some concepts focus on responses to story characters (e.g., identification, Cohen, 2001; para-social interaction; Hartmann & Goldhoorn, 2011) other concepts were introduced that cover the story experience as a whole. A large amount of research in this regard is based on narrative transportation (Gerrig, 1993; Green & Brock, 2000; 2002). Transportation is conceived as a holistic experience of allocating attention to a story, story-consistent cognitions and imagination, and strong feelings. Research suggests that the higher recipients' transportation the greater the enjoyment and the stronger the impact of stories on perceptions of oneself and the outside world (e.g., Green & Brock, 2000; Richter et al., 2014; Nielsen et al., 2018; van Laer et al., 2014; see Green et al., 2020). The extent to which recipients are transported into a story depends on the story, the recipient, the situation, and potential interactions between these three factors. Prior research on the antecedents of transportation demonstrated the relevance of the text itself (e.g., story structure and narrativity; Wang & Calder, 2006; Schreiner et al., 2018) as well as the influence of stable dispositions on the recipient's side, such as the need for affect (Appel et al., 2012; Appel & Richter, 2010). The research presented here deals with the third factor, the situation or circumstances under which story exposure occurs. More specifically, we were interested in the effect of screen size and the interaction between screen size and information encountered prior to story exposure.

Small Screens and Experiences Related to Narrative Transportation

The effect of screen size on the experience and effects of audiovisual stimuli (narrative or non-narrative) has attracted scholarly attention for many years now (e.g., Bellman et al., 2009; Dunaway & Soroka, 2021; Grabe et al., 1999). Although no study to date examined transportation as a dependent variable, hypotheses on related concepts were developed and tested empirically. In the very first issue of the present journal, Reeves and colleagues (1999) outlined several theoretical arguments as to why larger screens could and should enhance attention and arousal. Attention is a subcomponent of the holistic transportation experience (Green & Brock, 2000) and arousal is a key ingredient of strong emotions experienced in a state of high transportation (Green & Brock, 2000), at least when the arousal is elicited in congruence with the emotional responses implied by the story events (Appel et al., 2019; Tan, 1996).

Reeves and colleagues (1999) reasoned that larger screens have a greater likelihood of showing unfamiliar object sizes (e.g., displays of ants can be much larger than in real life) which increases attention. Moreover, larger screens have a greater likelihood of eliciting peripheral vision, as more of the information presented is outside the visual range of the fovea centralis. Peripheral vision is more sensitive to novelty and motion than foveal vision (Livingstone & Hubel, 1988), suggesting higher attention for larger screens. Another argument by Reeves and colleagues (1999) is that larger screens provide more searchable details than small screens (e.g., five trees close together are processed as different entities), increasing the likelihood for higher attention. Based on Kosslyn (1994), they further suggest that a larger-size image requires more attention because the resulting mental image is larger (or richer). With respect to arousal, Reeves and colleagues (1999) argue that there is a higher likelihood of novel displays on a large screen – such as eyes or other parts of the human body that are larger than in real life – that increase arousal. Moreover, objects on a larger screen dominate the visual field more and are more likely perceived as close or “looming”. As many audiovisual stimuli have the power to trigger emotional processing (aversive or appetitive), larger screens should therefore have a higher likelihood to elicit arousal.

In their experiment, Reeves and colleagues (1999) compared screen sizes of 56 inch, 13 inch, and 2 inch, and presented 60 short clips, each with a duration of six seconds. Physiological measures were used to quantify their main dependent variables attention and arousal. Data on heart rate deceleration in response to the onset of pictures and skin conductance responses suggested that the largest screen produced greater attention and arousal than both other screens. Importantly, no differences between the 13-inch and the 2-ich screen were found. In a recent study using psychophysiological measures (Dunaway & Soroka, 2021), participants were shown news segments on a laptop, presented on a computer screen with a smaller window (smartphone-size condition, 4.5 inch) or a larger window (13 inch). In this study, extending the results of Reeves and colleagues (1999), differences between the medium-sized and the small screen were found: Decreasing heart rate variability and skin conductance responses suggested that the larger screen elicited higher attention and arousal.

Much of the research on screen size and user experience is based on the concept of *presence*. Presence is defined as the extent to which participants experience a mediated world as non-mediated (Lombard & Ditton, 1997). Among the different forms of presence that have been distinguished after Lombard and Ditton's early work, spatial presence, "the superordinate feeling of being located within a virtual space", has received most attention (Cummings & Bailenson, 2016, p. 277). Transportation and spatial presence have a substantial theoretical overlap (Green et al., 2004; 2020) and some of the spatial presence self-report measures have item wordings similar to the Transportation Scale (Lee, 2004; Wissmath et al., 2009). Whereas transportation is a framework popular with research on linear, non-interactive media such as written texts or audiovisuals, the concept of presence is usually used in research on interactive media such as computer games and VR-environments (Skarbez et al., 2017). Relatedly, transportation requires stimuli that follow a narrative structure – this is not the case for presence.

Most studies suggest that bigger screens lead to higher spatial presence. Some early research on big versus even bigger screens found higher spatial presence for the larger-sized screens (e.g., IJsselsteijn et al., 2001; Lombard et al., 2000; Troscianko et al., 2012, Experiment 1), but other research found no such effect (e.g., Bracken & Pettey, 2007; Troscianko et al., 2012, Experiment 2). A meta-analytic summary by Cummings and Bailenson (2016) incorporated field of view as a predictor of spatial presence. Field of view was manipulated in primary studies with the help of blinders or display screen size of head-mounted displays or with TV or computer screens of different sizes. In these studies, screen resolution and viewing distance were held constant. The results suggest that field of view was a significant predictor of spatial presence, with a weighted r of .30, indicating that larger field of view (including larger screens) increases presence.

Newer research on smartphone screens appears to be in line with the superiority of larger screens. Kim and Sundar (2016) compared smartphones of two different sizes. Participants were randomly assigned to use a smartphone with a 5.3-inch screen or a smartphone with a 3.7-inch screen to browse through a website. At one point either an audiovisual ad or a text ad popped up and was watched by the participants. The audiovisual ads were pre-tested to contain contents with a low likelihood of eliciting strong emotions. Among a range of measures, one index (*being there-heuristic*) comprised five items of a presence scale (e.g., "I felt I could have interacted with the environment in the ad"). Participants who used the large-screen (5.3 inch) smartphone experienced a greater sense of 'being-there' than did those in the small-screen (3.7 inch) condition. The reported effect size was huge: The effect size of $\eta_p^2 = .48$ (which is equal to $d = 1.92$) surpassed, as a point of comparison, the difference in height between men and women ($d = 1.63$, Lippa, 2009).

Small Screens and Narrative Transportation: Cautionary Remarks on Existing Evidence

In sum, extant theory and research provide some evidence that screen size affects recipient experiences related to transportation. However, these results may not readily translate to the question of experiencing transportation when using a device with a small screen as compared to a medium-sized screen when watching a story. There are two reasons for our hesitation: The experimental setting of the primary studies and the stimuli used.

First, in the primary studies included in the meta-analysis on spatial presence (Cummings & Bailenson, 2016) and the studies on attention and arousal (Reeves et al., 1999; Dunaway & Soroka, 2021) viewing distance was held constant. Participants were prohibited to move closer to the smaller screen (an inclusion criterion for the meta-analysis by Cummings & Bailenson, 2016).[1] Whenever psychophysiological responses are measured (e.g., Dunaway & Soroka, 2021; Lombard et al., 2000; Reeves et al., 1999) restricting participant movement is required to obtain valid psychophysiological data. Thus, restricting movement can be a reasonable decision. That said, it runs counter to users' tendency to adapt screen distance (e.g., Bababekova et al., 2011; Charness et al., 2008; Paillé, 2015). Users faced with the situation that they are prohibited to move a handheld device closer to the eyes or move closer to a tiny screen could produce cognitions and emotions that

[1] Experimenters in the primary studies underlying the meta-analysis by Cummings and Bailenson (2016) tried to establish an identical distance in all conditions. Possibly, participants leaned forward and reduced their viewing distance in the respective smaller screen condition.

hinder and disrupt narrative transportation (Busselle & Bilandzic, 2008). Allowing participants to adjust distance could therefore enhance transportation when viewing on a small screen. As a consequence, a potential screen size effect would be reduced. Bellman and colleagues (2009) conducted an experiment on screen size in which participants could freely choose the distance to the screen. They included ads into sitcoms and measured delayed recall, attitude towards the ad/brand, and purchase intentions as their main dependent variables. These stimuli were presented on one of three different devices with screen sizes 35 inch, 10 inch, and 2 inch. The authors reasoned that if the screen size influenced attention and arousal (e.g., Reeves et al., 1999), this should have downstream effects on the dependent measures. Screen size, however, did neither affect their key dependent measures, nor the additional variables viewing comfort and viewing enjoyment. They found that participants located themselves much closer to the small screen than to both other screens and they found an association between viewing distance and the main dependent measures: The closer the participants located themselves (the wider the viewing angle), the more positive were the responses to the ads. These results underscore the possibility that screen size does not influence transportation whenever participants are free to choose the viewing distance (viewing angle) as they would in real-life viewing settings.

Second, the stimuli presented in prior studies had rarely been stories. The meta-analysis by Cummings and Bailenson (2016) summarized almost exclusively studies that involved computer games with low narrativity or some other, non-narrative environments in which participants had to navigate through. The narrative structure could reduce the advantage of larger screens. Theory on the superiority of larger screens has predominantly followed an early processing, bottom-up perspective to attention and related engagement (e.g., Reeves et al., 1999). Story processing, however, is characterized by rich representations of the events unfolding (Cutting, 2016; Graesser et al., 1997; Huff et al., 2014). Following a narrative constantly involves filling gaps (i.e., drawing inferences), building expectations, and processing incoming information based on these expectations (Busselle & Bilandzic, 2008; Gerrig, 1993; Huff et al., 2014; Zacks, 2013). Therefore, the perceptual differences of smaller versus larger screens could play a lesser role for narrative stimuli than for non-narrative stimuli (such as sets of short and distinct clips, Reeves et al., 1999; or environments to navigate through, Cummings & Bailenson, 2016). As a consequence, examining transportation into story worlds and comparing smartphone screen exposure to regular screen exposure could lead to effects that are smaller than observed in related prior research or even no effects at all.

Study Overview and Predictions

To date, there is no available study in which an audiovisual narrative was presented, screen size was manipulated, and transportation served as the dependent variable. Whereas indirect evidence concerning presence (Cummings & Bailenson, 2016) and psychophysiological measures (Reeves et al., 1999; Dunaway & Soroka, 2021) suggest a screen size effect on transportation, participants' regulation of viewing distance (Bellman et al., 2009) and the structure of stories and related processing (Buselle & Bilandzic, 2008) could diminish or even eliminate this effect.

If small screens (as compared to larger screens) impaired the experience of stories, this could have substantial implications on media influence in several fields, including health communication, marketing, or political communication. Moreover, a potential small screen effect could have substantial implications for research. Studies investigating story influence are regularly conducted online (e.g., Dessart, 2018; Johnson et al., 2020; Zwarun & Hall, 2012). In online studies it is often impossible to hold screen size constant, thus, participants' choice of screen could affect the results. Moreover, experimental procedures to manipulate transportation could be affected by screen size, and interaction effects could occur, pointing out the necessity to experimentally or statistically control for screen size in online studies.

The aim of three experiments was to examine the effect of presenting an audiovisual narrative on a small, smartphone-sized screen, as compared to a larger, monitor-sized screen, regarding recipients' experience of being transported into the narrative world. Based on our literature review that highlighted arguments for the effect of screen size as well as for the null hypothesis, we present two hypotheses for each research question – a null hypotheses and an alternative hypothesis – which will be tested with the help of classic null hypothesis significance testing and Bayes factor analyses. This notation will be applied to all research questions outlined. Our first pair of hypotheses addresses the main effect of smaller versus larger screens on narrative transportation.

$H1_0$: Smaller screens do not lead to less transportation into the story world than larger screens.

$H1_1$: Smaller screens lead to less transportation into the story world than larger screens.

We were particularly interested in the effect of screen size under circumstances that had previously yielded lower or higher transportation. Or, put differently, does screen size moderate the influence of other situational variables that are theorized to affect transportation? Regarding situations, several previous studies showed that information encountered prior to story exposure influenced the experience of the story (cf. Tukachinsky, 2014). Three types of information were investigated here, positive versus negative reviews (e.g., Tiede & Appel, 2020) were addressed in our first experiment, consistency of main character information (Busselle & Bilandzic, 2008) was manipulated in the second experiment, and prior information meant to facilitate understanding (Green et al., 2008) was focused on in our third experiment.

One of the factors that had shown a consistent effect on transportation is the presence of negative or positive reviews (Shedlosky-Shoemaker et al., 2011; Tiede & Appel, 2020). When participants read a negative review about a short film before exposure (as compared to a positive review or no review), transportation scores were reduced (e.g., Gebbers, et al., 2017; Isberner et al., 2019). This effect was attributed to negative (or positive) expectations elicited by the reviews that lead to confirmatory emotional responses and information processing (Tan, 1996; Tiede & Appel, 2020). Therefore, recipients' actual transportation into the story world is affected by the reviews.

Screen size and story reviews, however, could interact in their influence on transportation. Following prior theory and research on reviews and story experience (Tan, 1996; Tiede & Appel, 2020) recipients invest cognitive and emotional resources into processing media stimuli in expectation of gratifications. Positive reviews signal that more gratifications can be gained, hence, participants will allocate more attention leading to stronger transportation. This likely starts a self-reinforcing process: The more resources that have already been invested, the more resources will be invested in the future. Moreover, if a higher degree of transportation is already present at the story beginning, the evaluation of the stimulus will be more positive due to the less critical stance, and the prerequisite for more positive expectations will be fulfilled, leading to even higher transportation. Prior theory suggests that larger screens increase the likelihood for recipients to encounter unfamiliar object sizes, objects in peripheral vision, more details, novel depictions, and looming objects which increase attention and arousal elicited by these objects (Reeves et al., 1999). All these properties of large screens could increase the expectation-based effect of reviews, by providing substance to the negative or positive expectations. Other things equal, the self-reinforcing dynamics of expectations and gratifications should be stronger in a larger screen condition, given that an intense processing of details and high arousal (e.g., due to unfamiliar or looming objects) are more likely. In sum, the following expectations are put forward:

$H2_0$: Negative reviews do not lead to less transportation into the story world than positive reviews.

$H2_1$: Negative reviews lead to less transportation into the story world than positive reviews.

$H3_0$: There will be no interaction between screen size and review valence.

$H3_1$: An interaction between screen size and review valence will be observed.

A second kind of prior information discussed to influence transportation is information about the main protagonists. Busselle and Bilandzic (2008) refer to the character model in which recipients represent information about the protagonists (Graesser et al., 1997), such as their traits, motivations, and goals. Accordingly, recipients monitor continuously the protagonist and his/her intentions while following a story, and discontinuities in the protagonist's character or intentions lead to enhanced processing times, an indication for updating the character model (e.g., Rapp et al., 2001). We assume that such discontinuity effects will also occur if recipients built up a character model prior to watching a film, and this character model is inconsistent to information given in the film. If the character model is consistent with the information given in the film, no such discontinuities will emerge. We argue that discontinuities will enhance cognitive effort and reduce the fluency when watching the film (see for example Alter et al., 2007; Alter & Oppenheimer, 2009, for fluency-effects). Thus, readers will allocate cognitive resources to meta-cognitive processes in order to comply with the inconsistency instead of focusing on the construction of the mental model and of losing the focus of the narrative leads to less transportation into the film (cf. Busselle & Bilandzic, 2008). Given the increased attention allocated to a larger screen (Reeves et al., 1999), there should be a higher likelihood for inconsistencies to yield the consequences outlined above. Discontinuities may be overlooked when stimuli are presented on a small screen.

$H4_0$: Inconsistent main character information does not lead to less transportation into the story world than consistent main character information.

$H4_1$: Inconsistent main character information leads to less transportation into the story world than consistent main character information.

$H5_0$: There will be no interaction between screen size and consistency (fit) of character information.

$H5_1$: An interaction between screen size and consistency (fit) of character information will be observed.

A third kind of prior information discussed to affect transportation is knowledge about the film due to prior information about the story. Prior knowledge of the story provides readers with a mental model about the story, and readers are more familiar with the protagonists and the plot. When watching the film, the prior mental model may then enhance fluency of information processing, it leads to a deeper understanding, and, consequently, transportation is enhanced (Busselle & Bilandzic, 2008; Green et al., 2008). Indeed, Green and colleagues (2008, Study 1) found that participants who had read *Harry Potter and the Chamber of Secrets* before they watched the film were more transported into the film than participants who had not read the book before. This effect was replicated experimentally (Green et al., 2008, Study 2): Participants who read the story first and watched a film about the story second were more transported during the second exposure to the story than any other group in the experiment. The increased attention allocated to a larger screen (Reeves et al., 1999) should increase the fluency advantage in a situation in which a mental model of events unfolding could already be established.

$H6_0$: Prior information that improves understanding does not lead to higher transportation into the story world.

$H6_1$: Prior information that improves understanding leads to higher transportation into the story world.

$H7_0$: There will be no interaction between screen size and prior information.

$H7_1$: An interaction between screen size and prior information will be observed.

Methodological Considerations for the Experiments

Balancing Internal and External Validity

As outlined above, people's transportation into story worlds is a function of many factors. Beyond screen size, a factor that could affect transportation is the amount of distraction encountered outside of lab settings (e.g., Zwarum & Hall, 2012). Devices with smaller screens are more likely to be used outside of secluded spaces, therefore distraction could be larger when users watch stories on a small screen. Thus, in everyday use, screen size effects could be confounded with distraction effects (and potentially other factors that vary with mobile use). For research on screen size, this poses an inherent tension between internal validity and external validity. We tried to balance internal and external validity considerations by combining two lab experiments (context was held constant) and one online experiment (participants were free to choose where and when they participated in the study).

In all three experiments, participants were allowed to adapt their viewing distance, like they would outside the lab. Prior research indicates that users adjust the distance to a screen (by moving themselves or a handheld device) with the result that viewing distance is smaller for smaller screens than for larger screens (e.g., Bababekova et al., 2011; Charness et al., 2008; Paillé, 2015). This adaptation reduces small screen versus larger screen differences in terms of field of view, but these differences are still substantial: For smartphones, Paillé (2015) showed that participants chose an average mean distance of 33.8 cm from a 4inch smartphone screen (see also Ho et al., 2015; Bababekova, 2011), yielding a horizontal field of view of 14.9 degree (FOV = 2 x arctan (size/[2 x distance], 16:9 aspect ratio). For monitor screens participants choose an average mean distance of 68 cm (Charnass et al., 2008, see Jaschinsky, 2002, and Rempel et al., 2007 for similar estimates), yielding a field of view of 42.7 degree for a 24 inch screen. Thus, although users watch content presented on smaller screens from a closer distance, the field of view when using a smartphone-sized screen is still typically much smaller than the field of view when using a larger, monitor-sized screen.

A Priori Determination of Sample Size

We were not aware of any study that had investigated the effect of screen size on transportation. As a proxy, we assembled the effect sizes that had been found for presence as comparable experience measures. For presence, Lombard and colleagues (2000) identified an effect of η_p^2 = .071, and IJsselsteijn and and colleagues (2001) found an effect of η_p^2 = .093. These effect sizes are medium to large in size (Cohen, 1992), and therefore we calculated the needed sample size for a 2 (small vs. larger screen) x 2 (treatment) between participant ANOVA using f = 0.25. Given a type I error rate of .05 and a power of .80, we required 128 participants per experiment. For the three experiments somewhat higher sample sizes were aspired (Experiment 1: N = 140, Experiment 2: N = 135, Experiment 3: N = 200) as we assumed some drop-out due to technical issues or non-compliance with instructions, most notably for the online experiment (Experiment 3).

Statistical Analyses

We present additional information on the material and additional results pertaining to experience variables other than transportation in our online supplement. Our main hypothesis is complemented by a theoretically and practically relevant null hypothesis (i.e., small vs. larger screen size could not influence transportation for narrative stimuli). Thus, we added Bayes factor analyses to standard frequentist statistics (Dienes, 2014; Jarosz & Wiley, 2014; Konijn et al., 2015; Wagenmakers, 2007). The data and codes for all three experiments as well as the supplementary material are available online (https://osf.io/da827).

Experiment 1

Method

Design and Participants

To test our hypotheses, we conducted a 2 (screen size) x 2 (valence of review) between participants experiment in a research lab at the University of Würzburg, Germany. Transportation served as our main dependent variable. In addition, several additional measures were administered after the film was presented. Results on all ancillary measures are reported in the online supplement (benevolent sexism, hostile sexism, negative affect, positive affect, meaningful affect, and perceived corniness). In total 148 undergraduate students participated for course credit. We had to exclude 28 participants from the data analysis. One participant did not use the full screen mode to view the video, three participants reported that the audio was not okay, 20 participants knew the video before the study, one participant did not watch the entire video, and for three participants the data was incomplete due to a technical problem. Thus, the eligible sample consisted of 120 participants, 68 (56.7%) of them were female. The median age was 21 years with a range from 18 up to 53 years. Five students reported that German was not their mother tongue. We did not exclude them from the sample, as none of them had outlier values in any of the dependent variables and all of them indicated a language level of C1 (effective operational proficiency or advanced) or C2 (mastery or proficiency) in German. The participants were assigned to one of the four conditions: small screen - negative review ($n = 34$), small screen - positive review ($n = 28$), larger screen - negative review ($n = 27$), and larger screen - positive review ($n = 31$).

Materials

We showed the short film “Lieber Papa” [Dear Daddy] (CARE Norway, 2015), that was preceded by one of the two different review versions. The dependent measures were presented after the film.

The Film. The film Dear Daddy lasts 5:00 minutes in total and was intended to increase sensitivity with respect to sexism, sexual violence and harassment towards women by the producers. In the film, an unborn girl is telling her future story, which includes being exposed to sexist jokes at school, date rape, and, as she grows older, physical violence by her long-term partner. The story is told by a female narrator, with underlying slow-paced piano and strings music.

Reviews. Two versions of a review were developed based on a prior study (Gebbers et al., 2017). In the positive review version, the film was praised for its production quality and story, in the negative review version it was criticized for having low production value and an over-sentimental story. Both versions of the review were presented as supposed screenshots from a website and consisted of 107 words (positive version), or 106 words (negative version) respectively. Please refer to the online supplement for the exact wordings and the layout of the reviews. Participants were asked to write a brief, key points summary of the review to make sure that the reviews were read and understood (average number of words per participant: $M = 24.98$ words).

Transportation. Transportation ($M = 4.71$, $SD = 1.20$, Cronbach's alpha = .84) was assessed using the German short form of the Transportation Scale (Appel, Gnambs, Richter, & Green, 2015; Gebbers et al., 2017, e.g., "The video affected me emotionally"). All statements had to be answered on a seven-point response scale from 1 (not agree at all) to 7 (fully agree). Boxplots of the average score across the six items indicated five outliers that exceeded the whiskers, and we winsorized those values to one unit below the minimum value or one unit above the maximum value of the remaining scores (Tabachnick & Fidell, 2013).

Procedure

Participants were tested in groups up to a maximum of five persons, and participants were randomly assigned to either the computer screen condition or the smartphone screen condition depending on the time slot the participants selected from an online schedule. After giving their written informed consent, all participants were seated in front of a personal computer. They were randomly assigned to either the positive or the negative review condition. After reading the review, they were asked to write the brief summary. Next, they watched the short film Dear Daddy on the computer display in full screen mode (screen diagonal of 24 inch) or were asked to take a smartphone provided by the experimenter to watch the film (*Cubot* model *Rainbow*, screen diagonal of 5 inch). Participants used headphones in both conditions. Next, the participants filled in the dependent measures[2] on the computer. Finally, the participants answered socio-demographic questions and questions about the technical quality of the video and the sound before they were de-briefed. The experiment lasted on average around 15 minutes, all materials were presented using Unipark-software (www.unipark.com). All three experiments were approved by the research ethics committee of the Human Computer Media Institute at the University of Würzburg, Germany.

Results and Discussion

All null hypothesis significance tests were calculated with SPSS version 25, the type I error rate was set to .05. Partial eta-squared was calculated as a measure of the effect size and interpreted in terms of Cohen's (1992) taxonomy of effect sizes as small, medium or large corresponding to values of .01, .06, and .14, respectively. Additionally, we calculated Bayes-Factors using JASP Version 0.9.2. We used the default prior distributions that are implemented into JASP (see Rouder et al., 2012; Wagenmakers et al., 2018). We report BF_{10}, that is the likelihood of the data in favor of the H_1 compared to the null model (i.e. the model that includes only the grand mean and the subjects), and for interaction effects we additionally compared the full-effects model to the main-effects model. We checked the normal q-q-plots of transportation for each condition, and we found no severe deviations from a normal distribution.

[2] In addition to transportation, we assessed negative, positive, and meaningful affect, corniness, and sexism (please refer to online supplement for details).

Manipulation Check

To examine whether the valence of the review was perceived as being positive or negative we counted the positive and the negative statements in the summaries that the participants had written, and we calculated the difference between positive and negative statements. Two raters counted the statements independently from each other. The interrater reliability was high as the difference scores correlated significantly between raters, $r(118) = .91$, $p < .001$. Therefore, we used the scores of the first rater only. Participants who had read the positive review showed higher difference scores ($M = 1.32$, $SD = 1.03$) than participants who had read the negative review ($M = -2.15$, $SD = 1.00$), $t(118) = -18.80$, $p < .001$, $d = -3.43$, $BF_{10} = 5.24 \cdot 10^{33}$. Therefore, the data is much more likely under the H_1 (i.e., groups differ from each other) than under the H_0.

Effects of Screen Size and Review Valence

We calculated an ANOVA using transportation as the dependent variable and screen size and the valence of reviews as independent variables (see Table 1 for means and standard deviations and Figure 1 for a graphical display). We found no significant effect of screen size, $F(1, 116) = 1.86$, $p = .175$, $\eta_p^2 = .016$, $BF_{10} = .33$, and the Bayes factor was in favor of a null effect. The valence of the review had a significant effect on transportation, $F(1, 116) = 8.97$, $p = .003$, $\eta_p^2 = .072$, $BF_{10} = 7.58$, and the Bayes factor showed that the data is 7.58 times more likely under the H_1 than under H_0. The interaction effect was not significant, $F(1, 116) = 0.23$, $p = .634$, $\eta_p^2 = .002$. Comparing the full model to the main effects only model resulted in a Bayes factor of $BF_{10} = 0.23$, i.e. the data is approximately four times more likely under H_0 as it is under H_1. The results of the Bayes factor analyses are visualized in Figure 2.

In sum, our first experiment revealed the previously observed influence of review valence on transportation (e.g., Isberner et al., 2019; Tiede & Appel, 2020), but the small versus larger screen size manipulation appeared to have no effect. Note that descriptively the transportation scores in the small screen condition were higher than in the larger screen condition. Likewise, we found no indication that reviews had a distinct effect for videos presented on a small versus regular screen. To extend our database, our second experiment addressed the potential effect of screen size for a different film, and a different kind of information encountered prior to the story.

Experiment 2

Method

Design and Participants

We set up a 2 (screen size) x 2 (consistency of text) between participant experiment in the laboratory. Transportation served as the dependent variable. In addition, we examined fun and suspense as additional experience variables. Results on both variables are reported in the online supplement.

In total, 144 university students took part in the experiment for course credit. Five of them were excluded because they did not watch the entire video, resulting in a sample of 139 participants, 87 (62.6%) female. The age ranged from 18 up to 53 years with a median of 21 years. No student knew the video before, and all students indicated German as their mother tongue. Participants were assigned to one of the four conditions: iPod screen - consistent text (n = 34), iPod screen - inconsistent text (n = 38), computer screen - consistent text (n = 34), and computer screen - inconsistent text (n = 33).

Materials

The participants watched the film “Brüder” [engl.: Brothers] (Rübesam & Tsukernyk, 2013). One out of two versions of a short text that introduced the film to the participants was presented. The content of the introduction was either consistent or inconsistent with the film. Transportation was assessed after exposure.

The Film. A version of the short film “Brüder” was created (15:51 minutes, original length: 20:52 minutes) by eliminating parts of the beginning and the final credits. The film tells the story of two brothers. One brother, Mickey, robs an armored cash transport car that is driven by his brother Tom. Mickey threatens to kill a colleague of Tom and wants Tom to open the doors of the armored car. Tom refuses to open the doors because he is sure that his brother Mickey will not kill anybody. However, Mickey kills Tom’s colleague. Tom chases his brother and handles him over to the police.[3]

Main Character Introductions. Two texts were written that introduced the two main characters of the film to the participants before they watched the film. In the introduction the names of the two protagonists are mentioned, their relationship to each other in childhood and in adulthood is described, and their lives for the past three years are sketched. For Tom, further information about his low-paid job and his difficulties in developing stable relationships with women is provided. In the consistent version, the text was in accordance with the film’s content (134 words). In the inconsistent version (124 words) some facts contradicted the film: One of the brothers was named Max, the brothers were introduced being always close to each other from childhood to adulthood, and Tom earned a lot of money and lived in a stable relationship with his girlfriend. Thus, we assumed that the participants who read the inconsistent version of the text would be confused several times during the film, whenever the inconsistency of the information became apparent.

Transportation. To assess transportation eight items from the Transportation Scale by Green and Brock (2000) were used (*M* = 3.22, *SD* = 0.81, German version, Cronbach’s alpha = .81, see

[3] The editing process led to a suboptimal sound quality that was unintended. The sound was the same in both screen size conditions.

online supplement). All the statements in the questionnaires had to be answered on a five-point response scale from 1 (not agree at all) to 5 (fully agree). No outliers were detected using box-plots.

Procedure

Participants were tested in groups up to a maximum of five persons. All participants in one group were assigned either to the small or to the larger screen condition by random assignment depending on the time slot. Each participant was then randomly assigned to the consistent or inconsistent condition. All participants gave their written and informed consent to participate. Afterwards they read either the consistent or the inconsistent introduction of the main protagonists on the computer display. Then, the participants watched the film on a computer monitor in full screen mode (24-inch diagonal) or they watched the film on an iPod touch A1421 (4-inch diagonal) using headphones. Next, participants filled in sociodemographic questions and the transportation items. In the end, we provided two statements as a manipulation check (“The film and the text were consistent with each other”, “Reading the text has helped me to comprehend the story“). Both questions were provided with the answer options *yes*, *no* and *I am not sure*. The entire experiment lasted around thirty minutes (M = 38.6 minutes, SD = 4.9 minutes), all materials were presented using Unipark software.

Results and Discussion

We calculated frequentist and Bayes statistics (see Experiment 1). We checked the normal q-q-plots of transportation for each condition, and we found no severe deviations from a normal distribution.

Manipulation Check

To check if the participants recognized the inconsistencies in one version of the introduction, we calculated a cross-table between the two manipulation-check questions and the text version. Results indicated that the consistent text was judged as being more consistent than the inconsistent text, $\chi^2(2, N = 139) = 9.12$, $p = .010$, Cramer's $V = .26$ and it was judged as being more helpful to comprehend the film, $\chi^2(2, N = 139) = 12.13$, $p = .002$, Cramer's $V = .30$.

Main Analysis

We ran a between-participant ANOVA with screen size and consistency of the character introductions as independent variables and transportation as the dependent variable. Table 1 provides the descriptive results. The main effect of screen size was significantly in favor of the computer screen, $F(1, 135) = 4.54$, $p = .035$, $\eta_p^2 = .033$, $BF_{10} = 1.21$. This effect, however, was rather small and the Bayes factor was close to one and therefore inconclusive. The main effect of character introduction consistency was not significant, $F(1, 135) = 2.34$, $p = .128$, $\eta_p^2 = .017$, $BF_{10} = 0.49$, and the Bayes factor was inconclusive again. The interaction effect was not significant, $F(1, 135) = 1.13$, $p = .289$, $\eta_p^2 = .008$. Comparing the full model to the main effects only model resulted in a Bayes factor of $BF_{10} = 0.35$, i.e. inconclusive evidence concerning the interaction effect. Thus, whereas the frequentist approach showed a small but significant effect of the screen size on transportation, the Bayes factor did neither support an effect nor a non-effect of screen size.

The consistence of the main character introduction with the characters' behavior in the film had no influence on story experience and screen size did not moderate the influence of this context variable. We did, however, find a significant main effect of screen size on transportation. This effect was small in size and rendered inconclusive in the Bayes factor analysis. Given the somewhat contradictory results of the first two experiments a third experiment was conducted. This time we moved from a lab setting to an online experiment, reflecting the prevalent use of smartphones to watch videos at home, as part of everyday leisure activities or as a participant in an online study.

Experiment 3

Method

Design and participants

Our experiment was based on a 2 (screen size) x 2 (prior information) experimental design and was conducted online. Transportation served as our main dependent variable. We further assessed negative affect, positive affect, meaningful affect, and perceived corniness. Results on these measures are reported in the online supplement.

Participants were recruited by posting the link on social media and by students contacting friends and acquaintances. Participants were given the opportunity to take part in a lottery with a price of 100 Euro. In sum, 195 participants took part in the online experiment (five participants less than the aspired number of 200 participants). From those participants 165 completed the experiment. We excluded participants based on pre-defined exclusion criteria: Three participants reported that they did not understand the English text in the film, three participants failed a careless responding check item in the questionnaire, three participants were far too fast when reading the text (i.e., faster than 1000 words per minute), seven participants did not watch the entire video, four participants knew the video before the experiment, five participants reported that the video was not continuously streamed, one participant was not able to hear the spoken text in the video, six participants reported that they had been interrupted when watching the video, two participants reported that the video was not displayed fully on the screen or they did not hear the audio at all lacking a headphone, and two participants in the smartphone condition reported that they watched the video on a laptop. Among the remaining 129 participants, 90 (69.8%) were female; one participant (0.8%) did not report his/her gender. The youngest participant was 18 years old, the oldest 69 years (Md = 22 years). Most of the participants (83 or 64.3%) had obtained a school degree equivalent to high school, 30 (23.3%) had a college or university degree, 13 (10.1%) some lower level school degree, and three (2.3%) had not (yet) obtained a school degree. Three participants had another mother tongue than German, but as they provided no outlier scores in any of the analyses, they were retained in the sample. Participants were assigned to one of the four conditions: prior knowledge – computer (*n* = 45), prior knowledge – smartphone (*n* = 34), no prior knowledge – computer (*n* = 32), no prior knowledge – smartphone (*n* = 18).[4]

Materials

Our material consisted of the short film "Mind the Gap" (Flanagan, 2014), a text that introduced the film, and questions to assess transportation.

The film. The short film "Mind the Gap" (6:00 minutes) tells the true story of a woman named Margret, who visits a tube station in London every day to listen to her late husband – whose voice is used by the London underground to tell the passengers to mind the gap. Someday the voice of her husband is replaced by a female voice. The end of the video shows Margret sitting on her sofa and listening to the news on the TV. The speaker tells that Margret has written a letter to London Underground, and the company re-installs the voice of her husband at one tube station.

Character introduction. A short text (516 words) was prepared that tells the story of Margret. The text describes how Margret walks to the tube station every day, listening to the voice of her late husband. The text mentions the feelings of Margret listening to the voice, and her astonishment and grief when she realizes that the voice has been replaced by the voice of a female speaker. At this point, the written introduction ends in order not to spoil the story (even if spoilers do not seem to interfere with transportation, Johnson & Rosenbaum, 2015). The introduction was meant to facilitate the comprehension of the short film, but it did not tell the entire story.

[4] The exclusion of participants was almost equally distributed across the four conditions, $\chi(3, n = 165) = 4.36$, $p = .232$, Cramer's $V = .16$. The rather unequal sample sizes can be attributed to the randomization results of the software.

Transportation. Transportation (*M* = 5.11, *SD* = 1.06, Cronbach's alpha = .84) was assessed using the German short form of the Transportation Scale consisting of six items (Appel et al., 2015). The statements in the questionnaire had to be answered on a seven-point response scale ranging from 1 (not agree at all), 4 (partly) to 7 (fully agree). We winsorized values from three participants using the same procedure that we described in Experiment 1.

Procedure

The participants were asked to use a desktop computer or a laptop. First, they gave their written informed consent to take part in the online experiment. Then, they either read the description introducing the film (prior knowledge condition) or they read no text (no prior knowledge condition). After reading the introduction, the participants in the prior knowledge condition answered four multiple-choice questions about the story: (1) what is Margret doing each day, (2) what was the profession of her late husband, (3) what is especially important for her about the announcement at the tube station, (4) what is the name of her late husband. Each question had three answer options of which one was correct. We calculated the percentage of correct answers to check if participants had carefully read the text. Next, the participants filled in questions about their demographic data. Participants in the smartphone conditions were asked to switch to their smartphone now and to type in a link or scan a QR-code that was provided on the computer screen. Participants were asked to watch the upcoming video in full screen mode and to use a pair of headphones. After watching the film Mind the Gap, all participants filled in the transportation questionnaire on the computer screen with one embedded attention check-item (i.e., "select the option strongly agree here"). Finally, we asked some questions about the quality of the streaming, if the audio was understood, where they had watched the video, disturbances while taking part in the study, if they had known the video before, and which smartphone they had used. The study took approximately 15 minutes. Participants could take part in a lottery to win 100 Euros.

Results and Discussion

We calculated frequentist and Bayes statistics (see Experiment 1). We checked the normal q-q-plots of transportation for each condition, and we found no severe deviations from a normal distribution.

Preliminary analyses

To check if the participants in the prior knowledge conditions (n = 79) read the story carefully, we calculated the percentage of correct answers in the multiple-choice questions about the story introduction: 91% of the answers were correct. Thus, almost all participants were able to answer the questions correctly and we assumed that they read the text carefully. Further, we calculated the mean screen size of the smartphones (M = 5.06 inch, SD = 0.90 inch) for the 51 participants that provided data about their smartphone.

Transportation

We ran a 2 (smartphone vs. computer) x 2 (prior knowledge vs. no prior knowledge) between participants ANOVA using transportation as the dependent variable. Table 1 provides the descriptive statistics. The results of the ANOVA indicated no significant main effect for the device, $F(1, 125) = 0.13$, $p = .725$, $\eta_p^2 < .001$, $BF_{10} = 0.20$, and this result is also supported by the Bayes factor in favor of the null effect. We found no significant main effect for prior knowledge, $F(1, 125) = 2.77$, $p = .099$, $\eta_p^2 = .022$, $BF_{10} = 1.10$, and the Bayes factor was inconclusive. Finally, a non-significant interaction effect, $F(1, 125) = 1.07$, $p = .302$, $\eta_p^2 = .009$, was found, and comparing the full model to the main effects only model showed inconclusive evidence, $BF_{10} = .82$. To sum up, we found evidence against an effect of the device on transportation, inconclusive evidence for or against an effect of prior knowledge, and inconclusive evidence for or against an interaction effect of prior knowledge and device. In an additional analysis, we used the exact screen size of the participants' smartphones and correlated screen size with transportation for 50 participants in the small-screen condition who provided sufficient information about the smartphone that they had used during the study[5]. The correlation between screen size (*M* = 4.97 inch, *SD* = 0.61 inch) and transportation (*M* = 5.09, *SD* = 1.01) was not significant, $r(48) = .07$, $p = .626$, $BF_{10} = .20$. Note that the results reported in the results section remained virtually unchanged when we modified the exclusion criteria and included an additional 19 participants who had reported a disturbance when watching the video (e.g., their video stalled, somebody disrupted the participant).

Mini Meta-Analysis

Our main aim was to examine recipients' experience of being transported in an audiovisual story as a function of screen size (smartphone-sized vs. monitor-sized). The results of the three experiments differed remarkably, two studies (Experiment 1 and Experiment 3) yielded no effect of screen size whereas one study indicated – based on null hypothesis significance testing – that smaller screens could lead to less transportation (Experiment 2). Thus, we meta-analyzed the data to get a clearer overall picture (Goh et al., 2016). Across the three experiments, we compared transportation scores of the small screen group to the transportation scores of the larger screen group. The mini meta-analysis followed the recommendations by Goh and colleagues (2016). We used JASP version 0.11.1 to conduct the analysis. Cohen's *d* was chosen as the effect size metric and we used the fixed effects model. Across the three experiments, the difference between the screen size conditions amounted to $d = .074$, 95% CI [-.125; .274], $p = .465$. Taking together the results from our three experiments, this result indicates that recipients' story experience in terms of transportation is not reduced by watching stories on a small smartphone screen.

General Discussion

Smartphones have become a leading electronic device when users engage with audiovisual content (van Kessel et al., 2019). If watching videos on a smartphone-sized screen (as compared to a larger

[5] One participant was excluded because of an outlier value, i.e. a 9.7 inch smartphone display.

screen) reduced users' transportation into story worlds (Green et al., 2020; Dunaway & Soroka, 2021), this shallower experience could have an impact on the development of perspective taking, mentalizing, and empathy (e.g., Black & Barnes, 2015; Mar, 2018), targeted communication in the fields of health, marketing, and politics (e.g., Nielsen et al., 2018), and the communication of knowledge and culture in a society more generally (e.g., Gottschall, 2013; Pinker, 2007). In three experiments we examined the impact of presenting a story on a smartphone-sized screen, as compared to a larger-sized computer monitor screen. Throughout three primary studies and a mini meta-analysis we found no conclusive evidence that smaller screens impair story experience. Moreover, screen size did not affect the influence of pre-exposure information that had been used in the past to manipulate transportation (Green et al., 2008; Tiede & Appel, 2020).

These results inform theory on the influence of screen size on recipients' experiential states and related literatures in the social sciences and humanities (e.g., Chateau & Moure, 2016; Reeves et al., 1999) as well as theory on the antecedents of recipients' transportation into story worlds (Green & Brock, 2002; Green et al., 2020). Our results indicate that screen-size effects observed for non-narrative stimuli and dependent variables of presence, attention, and arousal (e.g., Cummings & Bailenson, 2016; Dunaway & Soroka, 2021; Kim & Sundar, 2016; Reeves et al., 1999) do not translate to the experience of transportation into audiovisual stories. These results underscore the importance of choosing appropriate stimuli, whenever researchers are interested in (technological) presentation modes and subsequent experiences. If stories are most relevant from a theoretical or practical point of view, stories should be used as stimulus material. Likewise, measures that are tailored to stories (such as narrative engagement or transportation) should be considered as dependent variables whenever these constructs fit the theoretical or practical insights to be gained. From an applied perspective the results indicate that media producers who wish to elicit narrative transportation should not be concerned with screen size effects. This notwithstanding, given our exclusive focus on transportation, we cannot rule out potential effects on presence, aesthetic emotions, or other aspects of story experience.

Our results are in support of the null hypothesis. As Rosenthal (1979) and others observed, results in support of the null hypothesis have a low likelihood to get published. This file-drawer problem is a challenge to the social sciences (Simmons et al., 2011; Vermeulen & Hartmann, 2015). In line with Konijn and colleagues (2015), we believe that the field of human technology interactions is full of practically relevant and theoretically plausible null hypotheses, yet there are strong incentives for showing that the newest technology leads to psychological dismay. In addition to academic incentives to report 'significant results' (Vermeulen & Hartmann, 2015), evidence for the detrimental influence of new media technologies is particularly welcomed by journalists, publishers, and the general public (Harcup & O'Neill, 2017). Our results suggesting that small screens do not lead to a more superficial engagement with stories complements recent research that found little support for the detrimental effects of newer technologies and applications (e.g., Appel et al., 2020; Orben et al., 2019).

Limitations and suggestions for future research

Empirical approaches in media psychology and research on screen size specifically are faced with the tension between internal validity and external validity. We combined two lab experiments and an experimental setting in the field to balance internal and external validity. In the lab experiments we controlled the devices used (including actual screen sizes) and situation variables. In Experiment 3, participants watched the videos on their own devices in the actual situation they were in. In this experiment, distractions could affect participants' transportation into story worlds equally in both conditions (such as the presence of others, Tal-Or, 2019), but the smartphone conditions could lead to even more distraction (e.g., higher likelihood of motor behavior by the mobile phone users). Given that Experiment 3 showed no difference between both screen size conditions (or interactions of this factor with prior information to facilitate understanding), we are confident that our results are not limited to tightly controlled settings. That said, we cannot rule out the notion that unobserved moderating variables of the situational context could influence the experience of stories followed on small versus larger screens (e.g., motor behavior, noise). Future research is encouraged to examine these possibilities.

We argue that the narrative form reduces mechanisms underlying the advantage of larger screens found in prior studies (e.g., Reeves et al., 1999). In the three experiments, three different stimulus films were presented in order to increase generalizability. Our research, however, had not been aimed at identifying moderating variables on the side of the films, such as movie genre or the use of film music. All three films told a story, it remains an open question whether films lower on narrativity (Abbott, 2002; van Laer et al., 2019) yielded larger effects.

To secure internal and external validity, we did not restrict the movement of participants. Participants were allowed to respond to our manipulation of screen size by shortening the distance from to their eyes to the screen. Thus, we did not attempt to control for the size of the visual stimuli on the retina as some earlier studies on screen size did (e.g., Dunaway & Soroka, 2021; Reeves et al., 1999; Troscianko et al., 2012; but see Bellman et al., 2009). This element of the experimental design was chosen on purpose. Outside the lab individuals adjust their distance to a device depending on screen size, particularly individuals adjust the distance to their handheld smartphone screen (e.g., Paileé, 2015). Restricting movement could have elicited cognitions (e.g., “why can’t I take the device and move it closer?”) and emotions (e.g., anger) that hinder transportation into story worlds, particularly in the small screen condition. Moreover, as outlined in the beginning of the methods section, moving the small screen closer to the eyes did not invalidate screen size differences regarding participants’ field of view. That said, future research is encouraged to examine screen size effects of narrative stimuli using a fixed distance approach, shedding light on a different range of user responses.

Further, we need to note that some of our results were inconclusive, therefore either larger samples are needed or sequential sampling procedures should be applied in future studies (Schönbrodt & Wagenmakers, 2018).

Finally, individual differences could serve as an additional moderator. Prior research found that stable individual differences, such as the need for affect (Appel & Richter, 2010) or narrative engageability (Bilandzic et al., 2019) predict how strongly recipients get transported into story worlds. Possibly, cohort and age effects could play a role when comparing the experience of stories shown on smaller versus larger screens. Unlike older generations, today's teenagers and young adults grew up with small screen storytelling (van Kessel, 2019), thereby reducing potential screen-size effects among this generation. Future research is encouraged to explore recipient variables to predict potential screen size effects, along with interactions with stimuli and situational predictors.

Conclusion

We examined recipients' experience of being transported into story worlds for audiovisual narratives shown on a smartphone screen as compared to a computer screen. Three experiments indicate that smaller screen size does not impair narrative transportation.

References

Abbott, H. P. (2002). *The Cambridge introduction to narrative.* Cambridge, UK: Cambridge University Press.

Appel, M., Gnambs, T., & Maio, G. (2012). A short measure of the need for affect. *Journal of Personality Assessment, 94*, 418-426. https://doi.org/10.1080/00223891.2012.666921

Appel, M., Gnambs, T., Richter, T., & Green, M. C. (2015). The transportation scale–short form (TS–SF). *Media Psychology, 18*, 243-266. https://doi.org/10.1080/15213269.2014.987400

Appel, M., Marker, C., & Gnambs, T. (2020). Are social media ruining our lives? A review of meta-analytic evidence. *Review of General Psychology, 24,* 60-74. https://doi.org/10.1177/1089268019880891

Appel, M., & Richter, T. (2010). Transportation and need for affect in narrative persuasion: A mediated moderation model. *Media Psychology, 13*, 101-135. https://doi.org/10.1080/15213261003799847

Appel, M., Schreiner, C., Isberner, M.-B., & Richter, T. (2019). The mediating role of event-congruent emotions in narrative persuasion. *Poetics, 77,* 101385. https://doi.org/ 10.1016/j.poetic.2019.101385

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*, 569-576. https://doi.org/10.1037/0096-3445.136.4.569

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*, 219-235. https://doi.org/10.1177/1088868309341564

Bababekova, Y., Rosenfield, M., Hue, J. E., & Huang, R. R. (2011). Font size and viewing distance of handheld smart phones. *Optometry and Vision Science*, *88*(7), 795-797. https://doi.org/ 10.1097/OPX.0b013e3182198792

Bellman, S., Schweda, A., & Varan, D. (2009). Viewing angle matters—Screen type does not. *Journal of Communication*, *59*(3), 609-634. https://doi.org/ 10.1111/j.1460-2466.2009.01441.x

Bilandzic, H., Sukalla, F., Schnell, C., Hastall, M. R., & Busselle, R. W. (2019). The Narrative Engageability Scale: A multidimensional trait measure for the propensity to become engaged in a story. *International Journal of Communication*, *13*, 801-832.

Black, J., & Barnes, J. L. (2015). Fiction and social cognition: The effect of viewing award-winning television dramas on theory of mind. *Psychology of Aesthetics, Creativity, and the Arts*, *9*, 423-429. https://doi.org/10.1037/aca0000031

Busselle, R. & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory, 18*, 255–280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Bracken, C. C., & Pettey, G. (2007). *It is REALLY a smaller (and smaller) world: Presence and small screens.* Paper presented at the 10th International Workshop on Presence, Barcelona, Spain.

CARE Norway (Producer). (2015). *Lieber Papa [Dear Daddy]* [short film]. Retrieved from https://www.youtube.com/watch?v=2yFJpTemg0M, English version https://www.youtube.com/watch?v=dP7OXDWof30

Charness, N., Dijkstra, K., Jastrzembski, T., Weaver, S., & Champion, M. (2008, September). Monitor viewing distance for younger and older workers. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 52, No. 19, pp. 1614-1617). Sage CA: Los Angeles, CA: SAGE Publications.

Chateau, D., & Moure, J. (2016). *Screens: From materiality to spectatorship–a historical and theoretical reassessment*. Amsterdam University Press.

Clark, D. (2018). More than 75% of worldwide video viewing is mobile. Retrieved from https://www.emarketer.com/newsroom/index.php/threequarters-video-viewing-mobile/

Cohen, Ja. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. https://doi.org/10.1037/0033-2909.112.1.155

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

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*, 272-309. https://doi.org/ 10.1080/15213269.2015.1015740

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781. Doi: 10.3389/fpsyg.2014.00781

Dessart, L. (2018). Do ads that tell a story always perform better? The role of character identification and character type in storytelling ads. *International Journal of Research in Marketing, 35*, 289-304. https://doi.org/10.1016/j.ijresmar.2017.12.009

Dill-Shackleford, K. E., Vinney, C., & Hopper-Losenicky, K. (2016). Connecting the dots between fantasy and reality: The social psychology of our engagement with fictional narrative and its functional value. *Social and Personality Psychology Compass*, *10*, 634-646. https://doi.org/ 10.1111/spc3.12274

Dunaway, J., & Soroka, S. (2021). Smartphone-size screens constrain cognitive access to video news stories. *Information, Communication & Society, 24*(1), 69-84. https://doi.org/ 10.1080/1369118X.2019.1631367

Flanagan, L. (Director, Writer, Producer). (2014). Mind the gap [short film]. Retrieved from https://www.youtube.com/watch?v=e56GqgMFp9A

Gebbers, T., De Wit, J. B. F., & Appel, M. (2017). Health communication - transportation into narrative worlds and the motivation to change health-related behavior. *International Journal of Communication, 11*, 4886-4906.

Gerrig, R.J. (1993). *Experiencing narrative worlds: On the psychological activities of reading.* New Haven, CT: Yale University Press.

Godard, J.-L. (1997, in Beugnet, M., 2013). Miniature pleasures: On watching films on an iPhone. In J. Geiger & K. Littau (Eds.), *Cinematicity in media history* (pp. 196-210). Edinburgh, UK: University of Edinburgh Press.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, *10*, 535-549. https://doi.org/10.1111/spc3.12267

Gottschall, J. (2013). *The storytelling animal: How stories make us human*. New York: Mariner Books.

Grabe, M. E., Lombard, M., Reich, R. D., Capanella Bracken, C., & Ditton, T. B. (1999). The role of screen size in viewer experiences of media content. *Visual Communication Quarterly, 6*(2), 4-9. https://doi.org/ 10.1080/15551399909363403

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163-189. https://doi.org/10.1146/annurev.psych.48.1.163

Green, M., Bilandzic, H., Fitzgerald, K., & Paravati, E. (2020). Narrative effects. In M.B. Oliver, A.A. Raney, & J. Bryant (Eds.). *Media effects. Advances in theory and research.* New York: Routledge. https://doi.org/ 10.4324/9780429491146-9

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. https://doi.org/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 & T.C. Brock (Eds.), *Narrative impact. Social and cognitive foundations* (pp. 315-342). Mahwah: Erlbaum.

Green, M. C., Kass, S., Carrey, J., Herzig, B., Feeney, R., & Sabini, J. (2008). Transportation across media: repeated exposure to print and film. *Media Psychology, 11*, 512-539. https://doi.org/10.1080/15213260802492000

Hamby, A., Brinberg, D., & Jaccard, J. (2018). A conceptual framework of narrative persuasion. *Journal of Media Psychology, 30*(3), 113-124. 10.1027/1864-1105/a000187

Harcup, T., & O'Neill, D. (2017). What is news? News values revisited (again). *Journalism Studies, 18*, 1470–1488. https://doi.org/ 10.1080/1461670X.2016.1150193

Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl revisited: Exploring viewers' experience of parasocial interaction. *Journal of Communication*, *61*, 1104-1121. https://doi.org/10.1111/j.1460-2466.2011.01595.x

Ho, J., Pointner, R., Shih, H. C., Lin, Y. C., Chen, H. Y., Tseng, W. L., & Chen, M. Y. (2015, August). Eyeprotector: Encouraging a healthy viewing distance when using smartphones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 77-85).

IJsselsteijn, W., de Ridder, H., Freeman, J., Avons, S. E., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments, 10*, 298-311. https://doi.org/10.1162/105474601300343621

Isberner, M. B., Richter, T., Schreiner, C., Eisenbach, Y., Sommer, C., & Appel, M. (2019). Empowering stories: Transportation into narratives with strong protagonists increases self-related control beliefs. *Discourse Processes, 56*(8), 575-598. https://doi.org/ 10.1080/0163853X.2018.1526032

Jarosz, A., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *Journal of Problem Solving, 7*(1), 1-9. https://doi.org/10.7771/1932-6246.1167

Jaschinksi, W. (2002). Proximity-fixation disparity curve and the preferred viewing distance at a visual display as an indicator of near vision fatigue. *Optometry and Vision Science, 79*, 158-169. https://doi.org/10.1097/00006324-200203000-00010

Johnson, B. K., & Rosenbaum, J. E. (2015). Spoiler alert: consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation. *Communication Research, 42*, 1068-1088. https://doi.org/10.1177/0093650214564051

Johnson, B. K., Udvardi, A., Eden, A., & Rosenbaum, J. E. (2020). Spoilers go bump in the night: Impacts of minor and major reveals on horror film enjoyment. *Journal of Media Psychology*. https://doi.org/10.1027/1864-1105/a000252

Konijn, E. A., van de Schoot, R., Winter, S. D., & Ferguson, C. J. (2015). Possible solution to publication bias through Bayesian statistics, including proper null hypothesis testing. *Communication Methods and Measures*, *9*, 280-302. https://doi.org/10.1080/19312458.2015.1096332

Kosslyn, S. (1994). *Image and brain.* MIT Press.

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, *14*, 27-50. https://doi.org/10.1111/j.1468-2885.2004.tb00302.x

Lippa, R. A. (2009). Sex differences in sex drive, sociosexuality, and height across 53 nations: Testing evolutionary and social structural theories. *Archives of Sexual Behavior, 38*, 631-651. https://doi.org/ 10.1007/s10508-007-9242-8

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740-749.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*, JCMC321. https://doi.org/ 10.1111/j.1083-6101.1997.tb00072.x

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

Mar, R. A. (2018). Stories and the promotion of social cognition. *Current Directions in Psychological Science*, *27*, 257–262. https://doi.org/10.1177/0963721417749654

Morris, B. S., Chrysochou, P., Christensen, J. D., Orquin, J. L., Barraza, J., Zak, P. J., & Mitkidis, P. (2019). Stories vs. facts: triggering emotion and action-taking on climate change. *Climatic Change*, *154*(1-2), 19-36. https://doi.org/10.1007/s10584-019-02425-6

Nielsen, J. H., Escalas, J. E., & Hoeffler, S. (2018). Mental simulation and category knowledge affect really new product evaluation through transportation. *Journal of Experimental Psychology: Applied, 24*, 145–158. https://doi.org/ 10.1037/xap0000169

Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences*, *20*, 618-628. https://doi.org/ 10.1016/j.tics.2016.06.002

Orben, A., Dienlin, T., & Przybylski, A. K. (2019). Social media's enduring effect on adolescent life satisfaction. *Proceedings of the National Academy of Sciences, 116*, 10226-10228. Doi: 10.1073/pnas.1902058116

Paillé, D. (2015). Impact of new digital technologies on posture. *Points de Vue - International Review of Ophthalmic Optics*, *72*, 22-30.

Pinker, S. (2007). *The stuff of thought.* New York, NY: Viking.

Rapp, D. N., Gerrig, R. J., & Prentice, D. A. (2001). Readers' trait-based models of characters in narrative comprehension. *Journal of Memory and Language, 45*, 737-750. https://doi.org/10.1006/jmla.2000.2789

Reeves, B., Lang, A., Kim, E. Y., & Tatar, D. (1999). The effects of screen size and message content on attention and arousal. *Media Psychology*, *1*(1), 49-67. https://doi.org/ 10.1207/s1532785xmep0101_4

Rempel, D.,Willms, K., Anshel, J., Jaschinski, W. & Sheedy, J. (2007). The effects of visual display distance on eye accommodation, head posture, and vision and neck symptoms. *Human Factors, 49*, 830-838. https://doi.org/10.1518/001872007X230208.

Richter, T., Appel, M., & Calio, F. (2014). Stories can influence the self-concept. *Social Influence*, *9*, 172-188. https://doi.org/10.1080/15534510.2013.799099

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638–641.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374. https://doi.org/10.1016/j.jmp.2012.08.001

Rübesam, D., & Tsukernyk, A. (Writers). (2013). Brüder [Brothers] [film]. In F. Baier & A. Maruhn (Producer). Ludwigsburg: Filmakademie Baden-Württemberg GmbH.

Ryan, M.-L. (2012). Narration in various media. In P. Hühn et al. (Eds.): *The living handbook of narratology*. Hamburg University.

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review, 25*, 128-142. https://doi.org/10.3758/s13423-017-1230-y

Schreiner, C., Appel, M., Isberner M.-B., & Richter, T. (2018). Argument strength and the persuasiveness of stories. *Discourse Processes, 55, 371–386*.

Shedlosky-Shoemaker, R., Costabile, K. A., DeLuca, H. K., & Arkin, R. M. (2011). The social experience of entertainment media: Effects of others' evaluations on our experience. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(3), 111-121. https://doi.org/ 10.1027/1864-1105/a000042

Simmons, J.P., Nelson, L.D., Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359–1366. https://doi.org/ 10.1177/0956797611417632

Skarbez, R., Brooks, Jr, F. P., & Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Computing Surveys (CSUR)*, *50*(6), 1-39. https://doi.org/ 10.1145/3134301

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6 ed.). Harlow: Pearson.

Tal-Or, N. (2019). The effects of co-viewers on the viewing experience. *Communication Theory*. Ahead-of-Print. https://doi.org/10.1093/ct/qtz012

Tan, E. S.-H. (1996). *Emotion and the structure of narrative film.* Mahwah, NJ: Erlbaum.

Tiede, K. & Appel, M. (2020). Reviews, expectations, and the experience of stories. *Media Psychology*, *23*(3), 365-390. https://doi.org/10.1080/15213269.2019.1602055

Tukachinsky, R. (2014). Experimental manipulation of psychological involvement with media. *Communication Methods and Measures*, *8*, 1–33. https://doi.org/ 10.1080/19312458.2013.873777

Troscianko, T., Meese, T. S., & Hinde, S. (2012). Perception while watching movies: Effects of physical screen size and scene type. *i-Perception, 3,* 414-425. https://doi.org/10.1068/i0475aap

Van Kessel, P., Toor, S. & Smith, A. (2019). A week in the life of popular YouTube channels. Pew Research Center. Retrieved from https://www.pewinternet.org/2019/07/25/a-week-in-the-life-of-popular-youtube-channels.

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. https://doi.org/10.1086/673383

Van Laer, T., Edson Escalas, J., Ludwig, S., & Van Den Hende, E. A. (2019). What happens in Vegas stays on TripAdvisor? A theory and technique to understand narrativity in consumer reviews. *Journal of Consumer Research*, *46*, 267-285. https://doi.org/ 10.1093/jcr/ucy067

Vermeulen, I., & Hartmann, T. (2015). Questionable research and publication practices in communication science. *Communication Methods and Measures, 9*(4), 189-192. Doi: 10.1080/19312458.2015.1096331

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p-values. *Psychonomic Bulletin & Review*, *14*, 779-804. https://doi.org/ 10.3758/bf03194105

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*, 58-76. https://doi.org/10.3758/s13423-017-1323-7

Wang, J., & Calder, B. J. (2006). Media transportation and advertising. *Journal of Consumer Research*, *33*, 151-162. https://doi.org/10.1086/506296

Wissmath, B., Weibel, D., & Groner, R. (2009). Dubbing or subtitling? Effects on spatial presence, transportation, flow, and enjoyment. *Journal of Media Psychology*, *21*, 114-125. https://doi.org/ 10.1027/1864-1105.21.3.114

Wojcieszak, M., Kim, N., & Igartua, J. J. (2020). How to enhance the effects of mediated intergroup contact? Evidence from four countries. *Mass Communication and Society, 23*(1), 71-106. https://doi.org/ 10.1080/15205436.2019.1630444

Zacks, J. M. (2013). Constructing event representations during film comprehension. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring cognition at the movies* (p. 227–243). Oxford University Press. https://doi.org/ 10.1093/acprof:oso/9780199862139.003.0012

Zwarun, L., & Hall, A. (2012). Narrative persuasion, transportation, and the role of need for cognition in online viewing of fantastical films. *Media Psychology, 15*, 327-355. https://doi.org/10.1080/15213269.2012.700592

Tables

Table 1

Transportation as a function of screen size and prior information

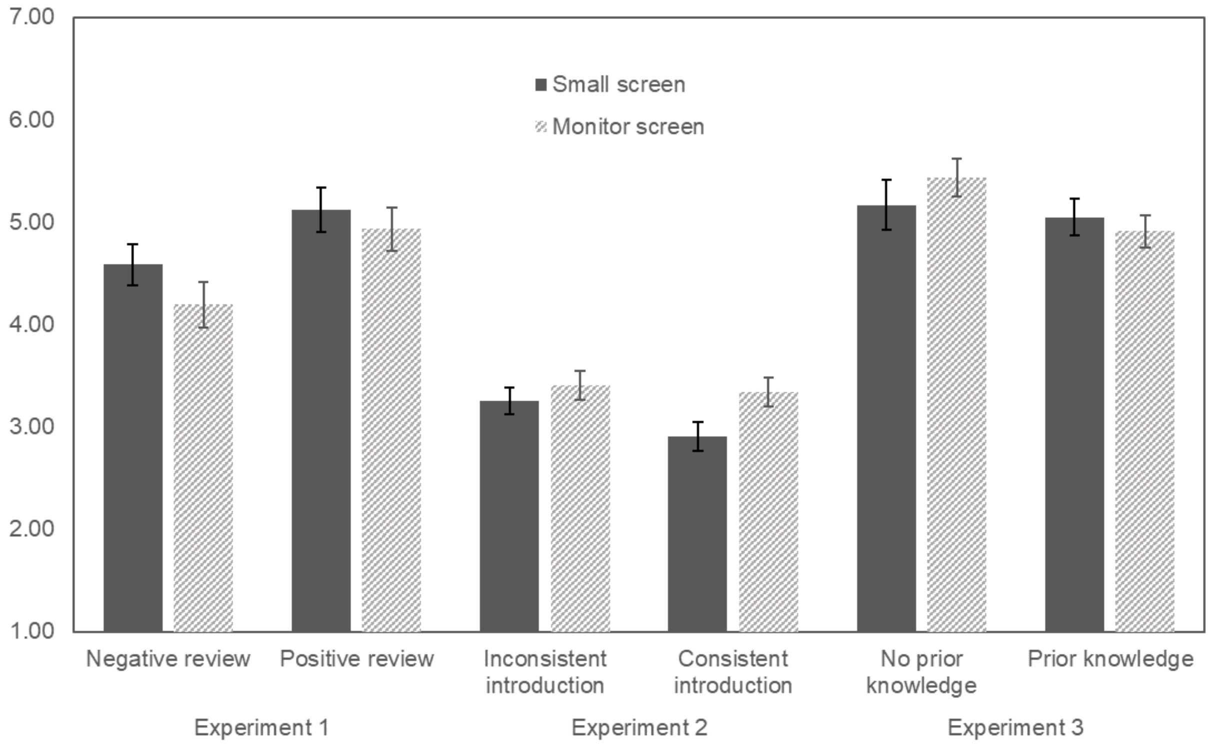
	Small Screen		Monitor Screen	
	M	*SD*	*M*	*SD*
Experiment 1				
Negative Review	4.58	1.28	4.19	1.32
Positive Review	5.12	0.90	4.93	1.08
Total	4.83	1.14	4.59	1.25
Experiment 2				
Inconsistent Character Introduction	3.26	0.77	3.40	0.98
Consistent Character Introduction	2.90	0.65	3.34	0.77
Total	3.09	0.73	3.37	0.87
Experiment 3				
Prior knowledge	5.04	1.12	4.91	1.21
No prior knowledge	5.17	0.75	5.44	0.84
Total	5.09	1.00	5.13	1.10

Notes. Experiment 1: 5-inch smartphone screen vs. 24-inch monitor screen; Experiment 2: 4-inch ipod screen vs. 24-inch monitor screen; Experiment 3: Own devices, smartphone ($M = 5.06$ inch) versus laptop or tablet screen. Transportation items ranged from 1 to 7 (Experiments 1 & 3), and from 1 to 5 (Experiment 2).

Figures

Figure 1

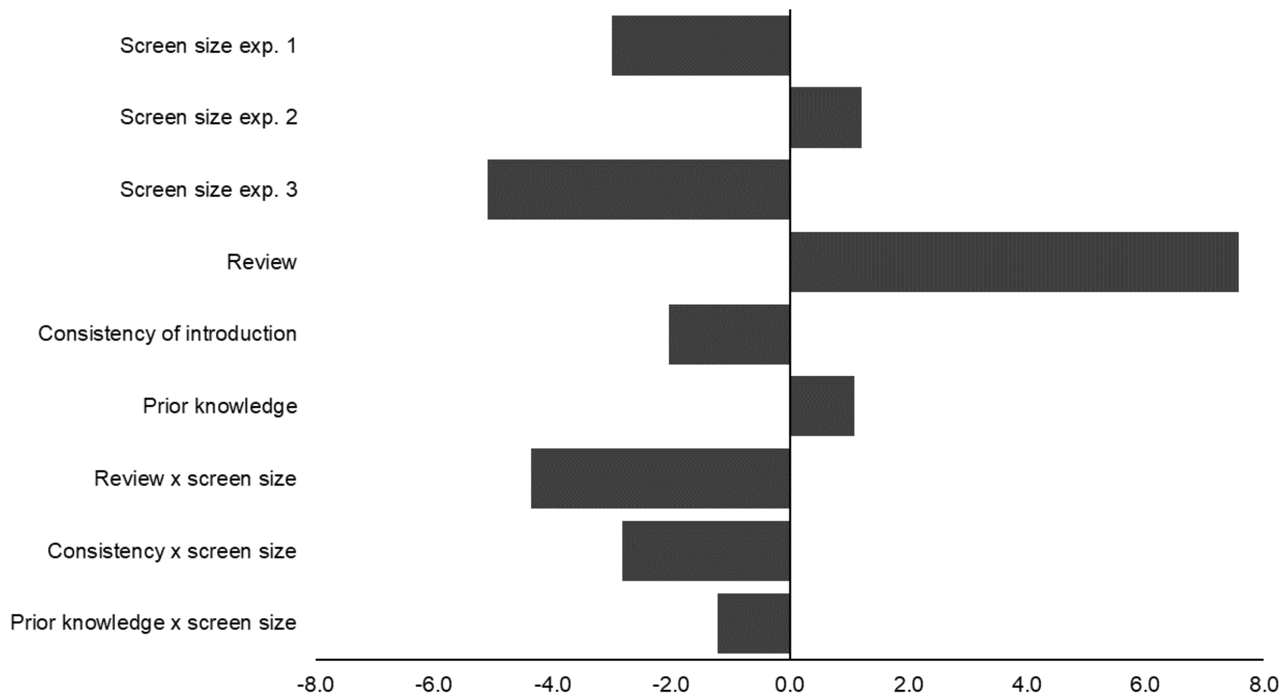
Transportation by experiment and condition.



Note. Values range from 1 to 7 in Experiments 1 and 3 and from 1 to 5 in Experiment 2. Error bars indicate standard errors of the mean.

Figure 2

Bayes-Factors (BF) across Experiments.



Note. For *BF* greater one BF_{10} is plotted, for *BF* smaller than one BF_{01} is plotted as a negative value. Positive values indicate evidence in favor of the H_1 (e.g., reduced transportation in the small screen condition), negative values indicate evidence in favor of the H_0 (e.g., transportation is not reduced in the small screen condition).