The Transportation Scale - Short Form (TS-SF)

Markus Appel
University of Koblenz-Landau

Timo Gnambs
Osnabrück University

Tobias Richter
University of Kassel

Melanie Green
State University of New York at Buffalo

Accepted for publication in *Media Psychology*

Author Note

Markus Appel, Department of Psychology, University of Koblenz-Landau, Germany. Timo Gnambs, Institute of Psychology, Osnabrück University, Germany. Tobias Richter, Department of Psychology, University of Kassel, Germany. Melanie Green, Department of Communication, State University of New York at Buffalo, USA.

This work was supported by grants from the Austrian Science Fund (FWF, I 996-G22) and the German Science Foundation (DFG, RI 1100/8-1) awarded to Markus Appel and Tobias Richter. We thank Barbara Malečkar for her support in conducting Study 1.

Correspondence concerning this article should be addressed to Markus Appel, Department of Psychology, University of Koblenz-Landau, Fortstr. 7, 78629 Landau, Germany. E-mail: appelm@uni-landau.de

Abstract

One popular concept to capture the experience of being immersed into the world of a story is the concept of *transportation*. This paper examines the factorial validity of the Transportation Scale (TS) and introduces a short and psychometrically sound alternative for the assessment of transportation. Exploratory bifactor modeling in Study 1 (N = 179) provided support for the hypothesized factor structure of the original TS with three facets and a general transportation factor. Based on these analyses, a six-item short version of the scale, the Transportation Scale- Short Form (TS-SF), was developed. Study 2 (N = 131 and N = 246) as well as Study 3 (N = 301) corroborated the short form's factorial structure, reliability, and validity.

Keywords: immersion, psychological state, persuasion, narrativity, narrative transportation

The Transportation Scale - Short Form (TS-SF)

Our lives are full of stories, such as a colleague's description of her last weekend, a newspaper article on the election in Chile, an episode of one's favorite TV series, or a true crime novel. Depending on the story, the situation, and person characteristics, individuals differ with regard to how deeply they get absorbed or *transported* into the world of a story in a particular instance. Green and Brock (2000) introduced the *Transportation Scale (TS)* as a method to quantify differences in the psychological states of being immersed in a narrative. Although the instrument proved quite influential in research on narrative persuasion (cf. van Laer et al., 2014), for many research designs the original scale turned out to be too long. Therefore, many authors administered various *ad hoc* shortened versions that only included subsets of items from the original TS (see below). When authors make individual decisions about which items to include, it becomes difficult to determine whether the construct was adequately measured and to compare results across studies. The aim of the present work was to examine the underlying factorial structure of the TS and to develop the *Transportation Scale-Short Form (TS-SF)*, a reliable and valid short instrument to measure transportation into narrative worlds.

Transportation into narrative worlds

In recent years, there has been a growing interest in the experience of (fictional) stories and the consequences thereof. Stories have been investigated in their capacity to influence recipients' factual knowledge (e.g., Dahlstrom, 2012; Marsh, Butler, & Umanath, 2012; Marsh, Meade, & Roediger, 2003), attitudes and beliefs (e.g., Appel & Richter, 2007; Gerrig & Prentice, 1991; Green & Brock, 2000), behavioral intentions and behavior (e.g., Appel & Mara, 2013), the self-concept (Djikic, Oatley, Zoeterman, & Peterson, 2009;

Richter, Appel, & Calio, 2014; Sestir & Green, 2010), and the theory of mind (e.g., Kidd & Castano, 2013; Mar & Oatley, 2008). The characteristic feature of following a story is that recipients often become immersed into the story world. This experience has been captured by the notion of *transportation* into narrative worlds (Gerrig, 1993; Green & Brock, 2000). When highly transported, "all mental systems and capacities become focused on the events occurring in the narrative" (Green & Brock, 2000, p. 701). Gerrig (1993) and Green and Brock (2000; 2002) described transportation as a rather broad concept that involves a co-activation of attention, imagery, and emotions. It incorporates the holistic experience of "being hooked" (Escalas et al., 2004) or lost in a story world (cf., Nell, 1988). Transportation, characterizing a psychological state, was found to be associated with trait measures of personality, such as the need for affect (Appel, Gnambs, & Maio, 2012; Appel & Richter, 2010) or trait absorption (Green & Brock, 2000). Moreover, transportation was also found to be related to a specific trait measure, "transportability" (e.g., Dal Cin, Zanna, & Fong, 2004; Mazzocco, Green, Sasota, & Jones, 2010). Transportability refers to individual differences in the tendency to become transported into a narrative world; transportability measures trait differences whereas transportation measures state differences (see Gnambs, Appel, Schreiner, Richter, & Isberner, 2014, for a latent state-trait analysis of transportation).

Transportation is a popular concept (cf. van Laer et al., 2014) but certainly not the only construct that addresses experiential states during media use. It is beyond the scope of this work to provide a thorough overview of the differences and similarities between transportation and alternative concepts. However, we wish to emphasize that unlike other concepts such as flow or presence (Csikszentmihalyi, 1992; Engeser, 2012; Wirth et al.,

2007), the concept of transportation is particularly focused on the experience of stories or narratives. It has been characterized as a holistic experience, an "integrative melding of attention, imagery, and feelings, focused on story events" (Green & Donahue, 2008, p. 241).

The Transportation Scale

The standard instrument to assess the experiential state of being transported into the narrative world is the Transportation Scale (TS, Green & Brock, 2000). It consists of 11 universal items and a varying number of items that address the readers' experience concerning the main characters of the specific story (four in the original version used by Green & Brock, 2000). These four items need to be adapted to include the character names (item wording: "I had a vivid image of [character name]"). Regarding the factorial structure of the TS, Green and Brock (2000) conducted an exploratory factor analysis that yielded a three-factor solution. The aspects represented by the three factors were labeled *cognitive* (items 1, 3, 4: "While I was reading the narrative, I could easily picture the events in it taking place"; "I could picture myself in the scene of the events described in the narrative"; "I was mentally involved in the narrative while reading it"), *emotional-affective* (items 5, 7, 11: "After finishing the narrative, I found it easy to put it out of my mind"; "The narrative affected me emotionally"; "The events in the narrative have changed my life"), and *visual imagery* (items 12-14: the four character name variables). Four items were unaccounted for by any of the three factors (items 2, 6, 8, 10: "While I was reading the narrative, activity going on in the room around me was on my mind"; "I wanted to learn how the narrative ended"; "I found myself thinking of ways the narrative could have turned out differently"; "The events in the narrative are relevant to my everyday life"). Green and Brock further

reported results of analyses that compared one-, two-, and three-factor solutions with the latter providing the best model fit (p. 704, no strict model comparison was presented). This initial analysis notwithstanding, the TS has been widely perceived as a measure for one transportation factor and in the large majority of studies, a sum or average score over all items was built. Despite its landmark position in research on narrative experience and narrative persuasion, the factorial structure of the TS has attracted little empirical attention since the original Green and Brock studies.

A further conundrum needs to be noted: Whereas almost all publications on transportation refer to Green and Brock (2000) and the TS as their measure of transportation, the number of items used in the measurement varied substantially. An inspection of the method sections of well-published papers in the field revealed that some researchers retained the original number of 15 items (e.g., Vaughn, Hesse, Petkova, & Trudeau, 2009; Wang & Calder, 2006), whereas others used shorter scales. Several numbers of items can be found, ranging from fourteen (Appel & Richter, 2010), thirteen (Richter et al., 2014), twelve (Dahlstrom, 2010; 2012), and ten items (Dunlop, Wakefield, & Kashima, 2010), to forms of only seven (Moyer-Guse & Nabi, 2010) or even three items (Escalas, 2004; 2007). Some of this variation can be explained as a result of different stimulus texts, because the transportation scale includes four items that address the visual imagery of four characters. Some stories used in the previous research might have included only three or even two characters; thus, a reduction of the number of visual imagery items might have been inevitable, and indeed, the scale was designed to be flexible in the number of imagery items. However, other variations go beyond simple adaptations to the number of characters in a story, resulting in differences that may have substantive effects.

Moreover, the length of the TS – fifteen items to measure one construct – might have appeared uneconomical, due to a need to include a number of different scales or because transportation was assessed as part of a repeated measures design. The proliferation of studies conducted over the Internet increases the need for short scales. Indeed, research on constructs tends to rise after a more economical measure is introduced (cf. Appel et al., 2012; Cacioppo, Petty, & Kao, 1984; Gnambs & Batinic, 2011).

Study overview

The first goal of the present paper was to examine the factorial validity of the TS. Theory and empirical practice suggest that the TS captures one latent factor that represents the holistic experience of being transported into the story world. This is not necessarily at odds with the results by Green and Brock (2000), who identified three factors which might be interpreted as three facets of transportation (cognition, emotion, imagery). Together with the general items (items that did not relate to any of the three factors) these items might also relate to a common transportation factor. This more complex model has never been examined statistically. In addition, by examining the relationships between all items and the general transportation factor, items can be identified that are not highly related to the general factor and might therefore be dispensable. This leads to our main goal, the development of a short form of the Transportation Scale. The short form will allow a more economical measurement of the general Transportation factor. This short form can be particularly helpful in studies where brief but psychometrically sound instruments are of key importance, such as studies with a repeated measures design and studies administered over the Internet or in naturalistic or field settings where brevity is essential.

Three studies were conducted. In Study 1 we analyzed the factorial structure of the TS. These analyses yielded information about common factor loadings of the items, which allowed the identification of items that were rather uninformative for the measurement of transportation. Study 2 was conducted to examine the psychometric properties of a short form of the TS, including the factorial structure, the reliability, its equivalence with the long form, and an analysis of the measurement invariance. In Study 3 an experimental design was chosen to gain additional evidence on the validity of the short form.

Study I: Factorial Validity of the Transportation Scale

Our first study examined the postulated factor structure of the TS. Following Green and Brock (2000), it was hypothesized that the instrument captures three facets of transportation: cognition, emotion, and imagination. We investigated to what degree the shared variance between items is attributable to a common factor representing the focal construct of transportation or to specific factors representing the three facets.

Method

Participants. We conducted a re-analysis of data that was published as part of unrelated research (Appel & Malečkar, 2012). Members of a British market research panel were invited to take part in the study which was conducted online. Among the original sample of 186 volunteers, seven had missing data on the TS, thus, our final sample comprised 179 participants (105 women) whose average age was 30.79 (*SD* = 7.05) years. About 35% of the participants had obtained a high school degree, 25% had a degree equivalent to one or two years of university education, and about 35% had completed university education.

Stimulus texts. The participants were randomly assigned to read one out of two stories. Story 1 was based on the short story *Murder at the Mall* (Nuland, 1994) which had been used in previous research on narratives and persuasion (e.g., Dal Cin, Zanna, & Fong, 2004; Green & Brock, 2000). This story takes place in a shopping mall. Two women and their children are introduced. Suddenly, a man approaches one of the children and kills the little girl. The attacker is identified to be a psychiatric patient with a history of violence. Story 2, *Christmas Carol*, was about a man with progressed cancer disease. He spends Christmas Eve with his family and numerous friends, knowing that it will be his last Christmas. Both texts originated from the same author and the same collection of stories (Nuland, 1994), and both stories were equivalent in length (969 words; control story: 951 words). We did not expect one story to be more transporting than the other. For the first story, different introductions regarding its truth status were administered, but for the purpose of this scale analysis, these groups were collapsed.

Measure. All participants were administered the Transportation Scale (TS; Green & Brock, 2000). The instrument included 15 items accompanied by seven-point response scales ranging from 1 (*not at all*) to 7 (*very much*). In the group that read the murder story, the scale resulted in a mean of 4.18 (*SD* = 1.01) and a coefficient alpha reliability of .89; in the group that read the Christmas story the respective mean and reliability were 4.09 (*SD* = 0.91) and coefficient alpha was .85.

Procedure. The entire questionnaire was administered as an un-proctored web-based test. Participants were invited by email and accessed the questionnaire from their home computers. Prior to answering the items of the TS, participants were randomly

assigned to read one of the two narratives (Story 1 was oversampled). There were 111 participants who read Story 1, and 68 participants who read Story 2.

Statistical Analyses. The factor structure of the TS was analyzed with a limited-information factor analysis for ordinal item responses (cf. Wirth & Edwards, 2007). Considering that the instrument is supposed to capture three facets of transportation (cognition, emotion, and imagination), we conducted an exploratory bifactor analysis (cf. Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Reise, 2012).[1] Bifactor models separate each item's variance into three parts: (a) variance attributable to a general factor common to all items, (b) variance attributable to a domain-specific factor representing each facet of transportation, and (c) residual variance unique to each item (see Figure 1). This bifactor specification allows for an examination of how much variance is accounted for by the hypothesized transportation factor and how much variance is due to the facet structure.

[1] Another obvious option to model the factor structure of the TS would have been second-order modeling. In this regard, it is important to note that the choice of bifactor modeling (BFM) versus second-order modeling (SOM) does not imply different theoretically views of the conceptualized measures. SOM and BFM are mathematically equivalent (cf. Chen, West, & Sousa, 2006; Chen et al., 2012); thus, each SOM can be formulated as an equivalent BFM. The major difference between both models is that the second-order (=general) factor in the SOM is assumed to exert indirect effects on each item, whereas the general factor in the BMF exerts direct effects. For the development of the short form of the TS, the general factor loadings are most informative. Accordingly, the BMF seems more appropriate because it allows examining the loadings of each factor on the general transportation factor. Moreover, the first-order disturbances in the SOM correspond to the domain-specific factors in the BFM. Thus, the BFM allows us to more directly investigate the amount of shared variance in each item due to the general and domain-specific factors (for an in-depth discussion of BFM in the area of personality research see Reise, 2012).

Thus, these analyses extracted four factors in total—one factor for each of the three facets and a general factor representing the common transportation construct. The resulting factor pattern was subjected to an oblique target rotation, which minimizes the difference between the estimated factor pattern and an *a priori* hypothesized target factor pattern (see Browne, 1972). The target factor pattern specified the bifactor model depicted in Figure 1: (a) The items hypothesized to form unique facets of transportation—cognition (Items 1, 3, and 4), emotion (Items 5, 7, and 11), and imagination (Items 12, 13, 14, and 15)—included loadings on their respective facet factors. (b) Cross-loadings of these items on the other facet factors were specified to be zero. (c) The remaining items not belonging to a facet were specified to have loadings of zero on the three facet factors. (d) All items loaded on the common factor. (e) The three facet factors were allowed to correlate. In contrast to confirmatory factor analyses, which force the loadings specified in (b) and (c) to be exactly zero, exploratory factor analyses with a target rotation estimate all cross-loadings such that they will be as close to zero as possible (Reise, 2012). Analyses that also estimate minor cross-loadings typically result in less biased parameter estimates than simple structure models (cf. Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013). Moreover, simulation studies have indicated that exploratory factor analyses with bifactor target rotation can recover true population parameters of bifactor loading patterns (Reise, Moore, & Maydeu-Olivares, 2011).

Results and Discussion

The results of the exploratory bifactor analyses in the two groups who read the different narratives are summarized in Table 1. Overall, the patterns of factor loadings in both groups approximated the hypothesized factor structure of the instrument well. Most

items had substantial loadings ($\lambda > .40$) on the common factor with a median loading of .61 in Story 1 and a median loading of .50 in Story 2. These results support the common practice of interpreting the sum score across all 15 items as indicative of a global transportation construct.

In addition, most items had also substantial loadings on their facet factors, indicating that the three facets explained unique variance beyond the general factor. Further information about the validity of the transportation scale can be gleaned from examining the variance of the observed scores attributable to the common factor. Factor saturation in a bifactor model is optimally ascertained by coefficient omega hierarchical ω_h (Zinbarg, Revelle, & Li, 2005), which is calculated as the sum of the factor loadings squared, divided by the model-implied variance of the scale scores. For the total item set, ω_h was .72 for the general factor in the group that read Story 1 and .73 in the group that read Story 2. By contrast, ω_h for the specific factors, i.e. the unique variance attributable to the three facets, was only .22 and .20, respectively. In sum, despite the presence of some multidimensionality in the 15 items, the majority of the observed score variance - about 70% - was attributable to a common transportation factor, and only about 20% of the variation was due to the three facet factors.

The general pattern of results corroborates the factorial validity of the Transportation Scale. However, the loading pattern in Table 1 also highlights the possibility of improving the measurement model through a revision of the TS. Several items had only marginal loadings on the common factor and therefore are rather uninformative for the measurement of transportation. For example, the loadings of Item 10 (“The events in the narrative are relevant to my everyday life”) were .11 and -.01 in the two groups. These

results strengthened our aim to develop a short form of the Transportation Scale. To this end, Study 2 was conducted.

Study 2: Development of the *Transportation Scale – Short Form (TS-SF)*

In order to develop a short form of the transportation scale, six items were selected based on the factor loadings obtained in Study 1, with the goal to retain the conceptual facets of the transportation construct. Two items were selected to represent the cognitive facet (Items 3 and 4 from the original scale) and another two items (Items 12 and 13) were selected to represent imagination (mental imagery). Among the items originally representing the emotion facet, only one item (Item 7) exhibited substantial loadings on the general factor (see Table 1, left column) and could therefore be included in a psychometrically sound short scale. We also included Item 6, which represented the common factor well and was part of those items that were not associated with a specific facet of transportation. The items of the short form are presented in the appendix. To explore the generalizability of our findings, Study 2 was based on two independent German-language samples (Study 2a and Study 2b). Study 2b employed a confirmatory bifactor model with cross-loadings of items constrained to zero, which yields a rather strict test of the hypothesized factorial structure of the short scale. In Study 2b we further examined the measurement invariance of the TS-SF for two different stories. We further present an analysis of factor congruence across all Study 1 and Study 2 samples.

Method

Samples. Study 2a was a re-analysis of a study conducted on the Internet (Appel & Richter, 2010, Study 2). The sample consisted of 133 volunteers; two participants had

missing data on the Transportation Scale. The remaining sample consisted of 131 participants (92 women) with a mean age of 30.63 years (*SD* = 12.22 years). Seventy-two percent of the volunteers had obtained a university entrance qualification or some college or university degree.

Study 2b was conducted as a paper-pencil experiment and included 251 Austrian volunteers. Five participants were excluded because of missing data on the Transportation Scale. Thus, the final sample of Study 2b consisted of 246 individuals (134 women) with a mean age of 25.48 years (*SD* = 4.42 years). Most participants (79%) had a university entrance qualification and 12% had a university degree.

Stimulus Texts. Study 2a presented one out of two stories about organ donation. The stories focused on a young man and two young women who met on campus and exchanged arguments about organ donation. In the beginning of Version 1, the young man was reminded of the dangers of car traffic; in the beginning of Version 2, the young man was hit by a car and died. The story in the latter version is about the events unfolding prior to his death. Both story versions were of equal length (421 words).

In Study 2b one out of two short fables was presented. The first fable dealt with a Native American chief and his two sons (451 words, *n* = 125 participants). The second fable featured a king, the king's son, and a fisherman (535 words, *n* = 121 participants).

Procedure and Measurement. In Study 2a and Study 2b, the participants were randomly assigned to read one of the respective stories. After reading the story, they worked on a German-language version of the TS (Appel & Richter, 2010). The German Transportation Scale was constructed following state-of-the-art standards in questionnaire language adaptation, involving a translation-back-translation procedure (Gudmundsson,

2009). In the present study, the German version included only three items on imagination because the stories used in Study 2 featured only three main characters. Thus, the TS consisted of 14 items with seven-point response scales ranging from 1 (*not at all*) to 7 (*very much*). In addition to the material outlined above, the participants worked on attitude items and individual difference measures not related to the analyses presented here.

Results and Discussion

Factorial Structure. In line with the previous study, we specified a bifactorial structure of the Transportation Scale – Short Form (TS-SF) that included a common factor for all six items and two domain-specific factors with two items each, representing the cognition and imagery facets. For the emotion facet no factor was specified because it included only one item. In Study 2a the factor structure was analyzed with a limited-information exploratory factor analysis with an oblique target rotation (Browne, 1972). For comparison, the respective results from the British sample in Study 1 are included as well. However, the results for the latter should be considered with caution because this sample was used to identify appropriate items for the short scale. Thus, it should exhibit a well-defined factor structure by design. The results of these analyses are summarized in Table 2. In both samples, the six items of the TS–SF exhibited substantial loadings on the general transportation factor. The shared variance between items was primarily captured by the transportation factor - ω_h was.71 and .84 - whereas the variance attributable to the facets was considerably smaller and fell at 6% and 21% (see Table 2).

In Study 2b we specified confirmatory models with a simple structure that constrained all cross-loadings of the items to zero. In the first sample the model resulted in

an acceptable fit[2] with $\chi^2(df = 8, N = 121) = 28.62$, $p < .001$, CFI = .96, TLI = .93, $\chi^2/df = 3.58$. In the second sample model fit was good with $\chi^2(df = 8, N = 125) = 31.78$, $p < .001$, CFI = .96, TLI = .93, $\chi^2/df = 3.97$. In both samples the items had strong loadings on the common transportation factor (median loadings: .58 and .65) which captured most of the shared variance between items, 74% to 81%. In contrast, the facets explained only a small amount of variance, 7% and 10% (see Table 2). Thus, using exploratory and confirmatory analyses, we identified the hypothesized transportation factor common to all items that explained most of the items' variances. This finding was obtained across different samples and stories.

Measurement Invariance. The comparison of mean transportation scores between two stories requires invariant factor structures (Millsap, 2011) – that is, the TS-SF should capture the same construct in a comparable manner for both stories. Measurement invariance was examined for the two stories in Study 2b as outlined by Millsap and Yun-Tein (2004). We specified a series of nested multi-group confirmatory factor models for ordinal item responses to examine invariance of factor loadings and item thresholds that are necessary for the comparison of latent means. In the first step, identical bifactor models were specified for both stories without placing any constraints on the estimated parameters. The satisfactory fit of this model (see Table 3) supported configural invariance indicating

[2] The fit of these models was evaluated in line with conventional criteria (cf. Chen, Curran, Bollen, Kirby, & Paxton, 2008; Marsh, Hau, & Wen, 2004) based on the comparative fit index (CFI) and Tucker-Lewis Index (TLI). Models with a CFI ≤ .95 or a TLI <.90 are considered "bad", and those with CFI > .95 and TLI > .90 as "good" fit.

similar factor structure of the TS-SF in both groups. Constraining the general factor loadings groups did not significantly impair model fit, $\Delta\chi^2(df = 4, N = 246) = 1.36$, $p = .85$, whereas putting additional constraints on the item thresholds resulted in a slightly worse model fit, $\Delta\chi^2(df = 33, N = 246) = 56.51$, $p = .01$. Partial scalar invariance was achieved after freeing the constraints on two thresholds of item 13 (see Table 3). Given the partial invariance latent means can be compared across groups by fixing the latent factor mean for one group to zero and estimating the mean for the other group. The latent group mean can then be compared on the basis of the *z*-statistic (Millsap & Yun Tein, 2004). The second story displayed a significantly larger latent mean, $z = 2.42$, $p = .02$, $d = .36$. A highly similar result also emerged for the comparison of the observed mean scores, $t(243) = 2.70$, $p = .01$, $d = .34$.

Reliability. Despite the reduction to less than half of the length of the original scale, the TS–SF did not exhibit an impaired measurement precision in comparison to the full scale (Table 4). Coefficient alpha reliabilities for the TS–SF ranged between .77 and .88 and thus, were comparable to the 14 item scale version.

Quantitative Equivalence. Descriptive analyses identified slightly higher mean scores for the TS–SF than for the TS. However, the respective effects were rather small (maximum $d = .34$), and the differences in five of the six groups were not significant (ps > .10, Table 4). The zero-order correlations between the short and long form of the Transportation Scale further underscored the equivalence of the two versions (see Table 3). The median cross-form correlation across all samples was .94. However, the TS–SF exhibited variances that were about two-thirds higher than the variances of the original TS.

Thus, the short form reflects individual differences in transportation more strongly than the TS.

Factor congruence. To determine the similarity of loadings patterns across the examined samples we calculated pairwise congruence coefficients (Tucker, 1951) for all Study 1 and Study 2 samples. These can be interpreted as correlations between loading coefficients of the same factor across different samples. Analyses by Lorenzo-Seva and Ten Berge (2006) suggest that congruence coefficients greater than .85 indicate fair similarity, whereas values of about .95 or greater imply factor loadings that can be considered equal in practice. The respective congruence coefficients for the general factor of the TS-SF (see Table 5) lay between .940 and .995 (*Mdn* = .98), indicating that across diverse stories and two language versions, the general factor reflecting the focal transportation construct can be regarded as equal.

In sum, our analyses suggest that the newly developed short form shows the expected factor structure with a strong, common transportation factor. An analysis of measurement invariance showed that the TS-SF measurement was equivalent across two different stories. The reliability of the short form is comparable to the reliability of the long form. Both versions are very highly correlated. An analysis of factor congruence across all samples of Study 1 and Study 2 provided additional evidence for equal factor loadings.

Study 3: On the construct validity of the Transportation Scale – Short Form (TS-SF)

After showing the expected factorial structure, a good reliability, and score comparability with the original long form of the scale, our aim was to shed light on another important aspect of the construct validity of TS-SF. To this end we compared the sensitivity

of the short form (TS-SF) and the long form (TS) with respect to a systematic manipulation of the stimulus material.

Method

Design. To examine the validity of the TS-SF, we conducted an experiment which involved three experimentally manipulated independent variables. Participants read a story and were randomly assigned to self-report their transportation into the story world with the help of the original Transportation Scale or the TS-SF (Independent variable 1: scale). Previous research indicated that transportation can be reduced by means of re-ordering the sections of a story (e.g., Wang & Calder, 2006). Thus, half of the participants read a story whose structure was manipulated beforehand (Independent variable 2: structure). To extend the generalizability of the results the participants were exposed to one out of two different stories (Independent variable 3: story). As a result, the experiment followed a 2 x 2 x 2 between-subjects design with random assignment.

Participants and Procedure. The experiment was conducted online and was programmed with EFS survey software. The participants were recruited through MTurk (Buhrmeister, Kwang, & Gosling, 2011) and received $0.30 for participation. After they read one out of four stories (see below) they completed either the six-item TS-SF or a 13-item long form of the Transportation Scale (only two character name items made sense for the chosen stories). The final part of the survey consisted of demographic measures.

The required sample size to identify effects of medium size was determined with G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). Because the study focused on the equivalence of the long and short scale versions, we decided to guarantee a substantial power of .95 and a type-I error probability of .05. These presumptions led to a minimum

sample size of 210. Because self-report surveys typically include a number of participants who adopt a careless response style (Mead & Craig, 2012), we oversampled and recruited 350 volunteers. Indeed, 41 individuals (11.71%) took less than 90 seconds to read the story which appeared to be a minimum time for thorough reading. One participant reported that his English was not as good as that of a native speaker. Seven additional participants gave a wrong answer when asked about the topic of the story. These participants were excluded from further analyses. The remaining sample consisted of 301 adults (149 women) between the ages of 18 and 73 (M = 37.14 years; SD = 12.94). All volunteers were from North America (297 from the US), 73.1% were Caucasian/White, 7.6% African American, 7.6% Asian American, 7.3% Latino/a, 3.0% indicated a mixed ethnic background, and 1.3% specified "other".

Stimulus Texts. The stimulus material was based on one out of two short stories. The first story was about a negative dating experience (1107 words), and the second story dealt with a pregnant woman and the relationship with her husband (834 words). Both stories were taken from a story website (onepagestories.com). For the purpose of this experiment we selected a common manipulation technique (e.g., Wang & Calder, 2006) and developed one version of each story in which the order of the events was altered. More specifically, the stories were divided into ten sections and the order of the sections was scrambled. In another version the story sections remained unchanged. Thus, four different texts were prepared, two different stories with either intact or scrambled section order.

Measures. The participants received the original Transportation Scale with two imagery items (as both stories included only two main characters) or the TS-SF. The survey

further included questions on demographics and a control question asked about the story's main topic.

Results and Discussion

We first inspected the reliabilities of both measures. For the long form of the dating story the intact and scrambled versions resulted in good reliabilities (Cronbach's alpha) of .77 and .78. The short form had reliabilities of .84 and .80. The pregnancy story showed reliabilities of .78 and .81 for the long form, whereas the short form resulted in reliabilities of .87 for the intact and scrambled story structure. For participants who received the original long form, we correlated the average score of all items and the average score of the six short form items, yielding correlations of .96 ($p < .001$) for the intact and scrambled version of the dating story. The respective correlations for the pregnancy story were .93 ($p < .001$) and .95 ($p < .001$).

To examine potential differences between both measures regarding the influence of story structure or story content, an ANOVA was conducted. The transportation score served as the criterion variable, and story (dating vs. pregnancy), structure (intact vs. scrambled), and scale (TS vs. TS-SF) were the experimental factors. We expected both the TS and the TS-SF to be sensitive to the story structure manipulation, with the scrambled stories eliciting lower transportation scores. The focal hypothesis referred to the interaction of story structure and scale. Given that comparable sensitivities to the experimental manipulation were assumed for TS and TS-SF, we did not expect to find a significant interaction effect.

We identified a significant main effect of the scale F (1, 293) = 28.90, $p < .001$, $\eta_p^2 = .09$, indicating that the long form (TS) yielded lower average scores ($M = 4.52$, $SD =$

0.92) than the short form (TS-SF, $M = 5.17$, $SD = 1.21$). Transportation scores for both stories differed, F (1, 293) = 33.55, $p < .001$, $\eta_p^2 = .10$; scores for the pregnancy story were higher ($M = 5.19$, $SD = 1.08$) than scores elicited by the date story ($M = 4.49$, $SD = 1.05$). The third main effect was significant as well, F (1, 293) = 5.78, $p = .017$, $\eta_p^2 = .02$: As expected, the intact stories elicited higher transportation than the scrambled stories ($M = 4.96$, $SD = 1.09$; $M = 4.71$, $SD = 1.14$). This main effect of story structure was qualified by a marginally significant two-way interaction, F (1, 293) = 3.75, $p = .054$, $\eta_p^2 = .01$), indicating that structural manipulation mattered for the date story (intact: $M = 4.73$, $SD = 1.01$; scrambled: $M = 4.25$, $SD = 1.05$) but not for the pregnancy story (intact: $M = 5.21$, $SD = 1.12$; scrambled: $M = 5.17$, $SD = 1.04$). Importantly, however, structure mattered irrespective of the scale used (short form or long form), as no significant two-way interaction between *structure* and *scale* was found, F (1, 293) = 0.46, $p = .497$, $\eta_p^2 = .00$. Moreover, the difference between the date-story and the pregnancy-story was present regardless of which scale was used, as indicated by a non-significant two-way interaction between *story* and *scale*, F (1, 293) = 0.03, $p = .866$, $\eta_p^2 = .00$. The three-way interaction did not reach significance, F (1, 293) = 3.00, $p = .085$, $\eta_p^2 = .01$.

In sum, this experimental study points at the validity of the Transportation short form. The TS-SF responded as sensitively as the TS towards a manipulation of story structure or differences between stories.

General Discussion

In recent years, there has been a proliferation of studies in psychology, communication science, and marketing that deal with the experience and effects of stories (cf. van Laer et al., 2014). A construct that has received widespread attention regarding the

process of experiencing narrative worlds is *transportation* (Gerrig, 2000; Green & Brock, 2002; van Laer et al., 2014). This experiential state has been characterized as an "integrative melding of attention, imagery, and feelings, focused on story events" (Green & Donahue, 2008, p. 241). The standard instrument to assess transportation is the Transportation Scale (Green & Brock, 2000) which includes 15 items and is typically treated as a one-dimensional measure (cf. van Laer et al., 2014). The first aim of this research was to examine the factorial structure of the Transportation Scale and to identify items that constitute a reliable and valid short form of the Transportation Scale. Our analysis of the underlying factor structure yielded some support for a three facets account (cognition, emotion, affect, Green & Brock, 2000), but most dominant was a general transportation factor. However, not all items showed substantial loadings on this factor, indicating the possibility to shorten the Transportation Scale without losing important information. A six-item version was developed, the TS-SF, that yielded the expected one-dimensional factor structure and good reliability scores. Further, it was shown to be as sensitive as the long form with regard to systematic textual variations, speaking for its validity. The TS-SF is available in English and German.

The current studies make both theoretical and methodological contributions. Theoretically, these analyses provide empirical support for the conceptual definition of transportation as an *integrative melding* of affect, cognition, and imagery. Although the individual facets of transportation make some unique contribution to transportation effects, the experience is well-captured by the general factor that combines these elements. Furthermore, the current paper makes use of a novel statistical technique, bifactor analysis, that may be a valuable tool for other communication studies.

These findings suggest that the individual difference version of the Transportation Scale, a measure of transportability (one's tendency to become immersed in narratives more generally), might also be effectively measured with a short scale. Mazzocco and Green (2011) proposed a four-item short form of the transportability measure, but their analysis was limited to showing a high correlation between the subset of items and the full scale. Additional research to develop a short transportability scale based on the current Transportation Scale – Short Form would be a useful future direction.

The TS-SF includes two imagery items which focus on the main characters of the story. We recommend that if there are more than two characters in a story that researchers use the two most important characters, which should be sufficient to capture variation in imagery across readers. For stories with only one character, we suggest including the item "I had a vivid mental image of the story setting(s)" as a second imagery item.

A limitation of the present research, as with much of the research in the field, is that the TS-SF is a self-report scale that asks for describing a past experience (even if the experience often took place only minutes or even seconds ago). As such, the responses are prone to potential biases in memory and judgment. Clearly more research is needed to examine the experience of stories, including affective, imaginative, and cognitive responses, as they happen (preferably without disrupting the experience itself).

As another limitation, we need to note that – like Green and Brock's (2000) original research – our studies focused on written texts and reading. The transportation experience, however, has been applied to the experience of audiovisual material as well (e.g., Williams, Green, Kohler, Allison, & Houston, 2010; Green et al., 2008). We have no reason to doubt that the TS-SF can be easily adapted to audiovisual texts with simple wording changes

(such as replacing “reading” with “viewing”, “watching”, or “listening to”). Nonetheless, future research in this direction seems warranted.

In particular, one important difference between audiovisual material and text is that the mental imagery is to some extent provided to viewers of video or other visual content. The question then arises whether it is sensible to include the “vivid image” items for this type of stimulus. To provide some initial guidance for this question, we re-examined the data from the video condition in Study 2 of Williams and colleagues (2010). The Cronbach’s alpha for the shortened scale was approximately .10 higher for the version with the imagery items included. Furthermore, although responses to the imagery items tended to be above the scale midpoint (e.g., 3 or higher), there was substantial variance within this range, suggesting that these items are tapping differences in participants’ experience. Therefore, we recommend retaining the imagery items for video stimuli.

An additional potential limitation is that because the short form contains fewer items assessing each component of transportation, researchers may be more limited in their ability to explore the relative contribution of emotion versus imagery, for example. If these types of analyses are the main focus of a particular set of studies, researchers can use the original scale or use other measures designed for that purpose, such as Busselle and Bilandzic’s (2009) narrative engagement scale. However, many current studies examine transportation as a unitary construct, and the current short form will have the significant advantage of allowing greater consistency across studies. Researchers will be able to have an economical measure without relying on intuition or individual decisions about which items to include or exclude.

Finally, there is an important caveat to our work that needs to be considered: No single set of studies can "validate" a scale (Cronbach & Meehl, 1955). Rather the validity of a scale is the product of the experiences a number of researchers gather over time. Thus, this work is not an end, but a starting point to ascertain the validity of the TS-SF. Particular attention might be given to the measurement invariance of our scale. The Transportation experience is conceived to be a function of the recipient, the situation, the mediated stimulus, and related interactions. Unlike trait measures, state measures in the field of media psychology need to capture differences in media stimuli. In Study 3 the TS-SF was indeed found to be sensitive to theory-guided variations of stimulus texts. Study 2 provided first evidence of measurement invariance (Millsapp, 2011), i.e., the TS-SF captured transportation in a comparable manner for two German language stories. Future research is encouraged to add to this evidence by focusing on the English language scale and by including other stimuli in their tests of measurement invariance. The future role of the TS-SF will likely be that of a research instrument in academics and market research. Mean scores in response to different media stimuli will be compared, and the scores will be used as dependent variables, predictors, indicators of mediation or moderation, or as covariates. We believe that for these purposes, the psychometric properties are sufficiently established. However, at this point we are cautious about recommending the TS-SF as a measure for individual diagnostics, e.g., within the framework of educational counseling or in clinical research (e.g., as an indicator of alexithymia).

In sum, the research presented here shows that the TS-SF can be a helpful research tool to assess the processing of narrative stimuli. Given the proliferation of narratives in our everyday lives in connection with the key role of transportation to explain story effects on

knowledge, attitudes and beliefs, the self, and theory of mind, the TS-SF is expected to become a highly popular research instrument. Its usage will likely include academia as well as various fields of applied communication and market research.

References

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

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

Appel, M. & Mara, M. (2013). The persuasive influence of a fictional character's trustworthiness. *Journal of Communication, 63*, 912-932. doi: 10.1111/jcom.12053

Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology, 10,* 113-134. doi: 10.1080/15213260701301194

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

Browne, M. W. (1972). Oblique rotation to a partially specified target. *British Journal of Mathematical and Statistical Psychology, 25*, 207-212. doi: 10.1111/j.2044-8317.1972.tb00492.x

Buhrmeister, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 1-5. doi: 10.1177/1745691610393980

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

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the

bifactor model to other approaches. *Journal of Personality, 80*, 219-251. doi: 10.1111/j.1467-6494.2011.00739.x

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of Quality of Life. *Multivariate Behavioral Research, 41*, 189–225. doi:10.1207/s15327906mbr4102_5

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302. doi:10.1037/h0040957

Csikszentmihalyi, M. (1992). *Flow: The psychology of happiness*. New York, NY: Harper & Row.

Dahlstrom, M. F. (2010). The role of causality in information acceptance in narratives: An example from science communication. *Communication Research, 37*, 857-875. doi: 10.1177/0093650210362683

Dahlstrom, M. F. (2012). The persuasive influence of narrative causality: Psychological mechanism, strength in overcoming resistance, and persistence over time. *Media Psychology, 15*, 303-326. doi: 10.1080/15213269.2012.702604

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

Djikic, M., Oatley, K., Zoeterman, S., & Peterson, J. B. (2009). On being moved by art: How reading fiction transforms the self. *Creativity Research Journal, 21*, 24-29. doi: 10.1080/10400410802633392

Dunlop, S. M., Wakefield, M., & Kashima, Y. (2010). Pathways to persuasion: Cognitive and experiential responses to health promoting mass media messages. *Communication Research, 37*, 133-164. doi: 10.1177/0093650209351912

Engeser, S. (2012). *Advances in flow research*. New York, NY: Springer.

Escalas, J. E. (2004). Imagine yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising, 33,* 37-48. doi: 10.1080/00913367.2004.10639163

Escalas, J. E. (2007). Self-referencing and persuasion: Narrative transportation versus analytical elaboration. *Journal of Consumer Research, 33*, 421-429. doi: 10.1086/510216

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. doi: 10.3758/BRM.41.4.1149

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

Gerrig, R. J., & Prentice, D. A. (1991). The representation of fictional information. *Psychological Science, 2*, 336-340. doi: 10.1111/j.1467-9280.1991.tb00162.x

Gnambs, T., Appel, M., Schreiner, C., Richter, T., & Isberner, M.-B. (2014). Experiencing narrative worlds: A latent state-trait analysis. *Personality and Individual Differences, 69*, 187-192.

Gnambs, T., & Batinic, B. (2011). Evaluation of measurement precision with Rasch-type models: The case of the short Generalized Opinion leadership Scale. *Personality and Individual Differences, 50*, 53-58. doi: 10.1016/j.paid.2010.08.021

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

Green, M.C., & Donahue, J. (2008). Transportation in narrative processing. In K. Markman, W. Klein, & J. Suhr (Eds.), *The handbook of imagination and mental simulation* (pp. 241–256). New York, NY: Psychology Press.

Green, M.C., Kass, S., Carrey, J., Feeney, R., Herzig, B., & Sabini, J. (2008). Transportation across media: Print versus film comparisons. *Media Psychology, 11*, 512-539.

Gudmundsson, E. (2009). Guidelines for translating and adapting psychological instruments. *Nordic Psychology, 61*, 29-45. doi: 10.1027/1901-2276.61.2.29

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, *342*, 377-380. doi: 10.1126/science.1239918

Lorenzo-Seva, U., & Ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, *2*, 57-64. doi:10.1027/1614-1881.2.2.57

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *13*, 173-192. doi: 10.1111/j.1745-6924.2008.00073.x

Marsh, E. J., Butler, A. C., & Umanath, S. (2012). Using fictional sources in the classroom: Applications from cognitive psychology. *Educational Psychology Review, 24*, 449-469. doi: 10.1007/s10648-012-9204-0

Marsh, E. J., Meade, M. L., & Roediger, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language, 49*, 519–536. doi: 10.1016/S0749-596X(03)00092-5

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341. doi: 10.1207/s15328007sem1103_2

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, *18*, 257–284. doi: 10.1037/a0032773

Mazzocco, P. M., Green, M. C., Sasota, J. A, & Jones, N. W. (2010). This story is not for everyone: Transportability and narrative persuasion. *Social Psychology and Personality Science, 1,* 361-368. doi: 10.1177/1948550610376600

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*, 437–455. doi: 10.1037/a0028085

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39,* 479-515. doi:10.1207/S15327906MBR3903_4

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

Nuland, S. (1994). *How we die: Reflections on life's final chapter.* New York, NY: Knopf.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667-696. doi: 10.1080/00273171.2012.715555

Reise, S. P., Moore, T., & Maydeu-Olivares, A. (2011). Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. *Educational and Psychological Measurement, 71,* 684-711. doi: 10.1177/0013164410378690

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

Sestir, M., & Green, M. C. (2010). You are who you watch: Identification and transportation effects on temporary self-concept. *Social Influence, 5,* 272-288. doi: 10.1080/15534510.2010.490672

Tucker, L. R. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.

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

Vaughn, L. A., Hesse, S. J., Petkova, Z., & Trudeau, L. (2009). "The story is right on": The impact of regulatory fit on narrative engagement and persuasion. *European Journal of Social Psychology, 39*, 447-456. doi: 10.1002/ejsp.570

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

Williams, J. H., Green, M. C., Kohler, C., Allison, J. J., & Houston, T. K. (2011). Stories to communicate risks about tobacco: Development of a brief scale to measure transportation into a video story. *Health Education Journal, 70*, 184-191. doi: 10.1177/0017896910373171

Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: Current approaches and future directions. *Psychological Methods, 12*, 58-79. doi: 10.1037/1082-989X.12.1.58

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., ... & Jäncke, P. (2007). A process model of the formation of spatial presence experiences. *Media Psychology, 9*, 493-525. doi: 10.1080/15213260701283079

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω_H: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*, 123-133.doi: 10.1007/s11336-003-0974-7

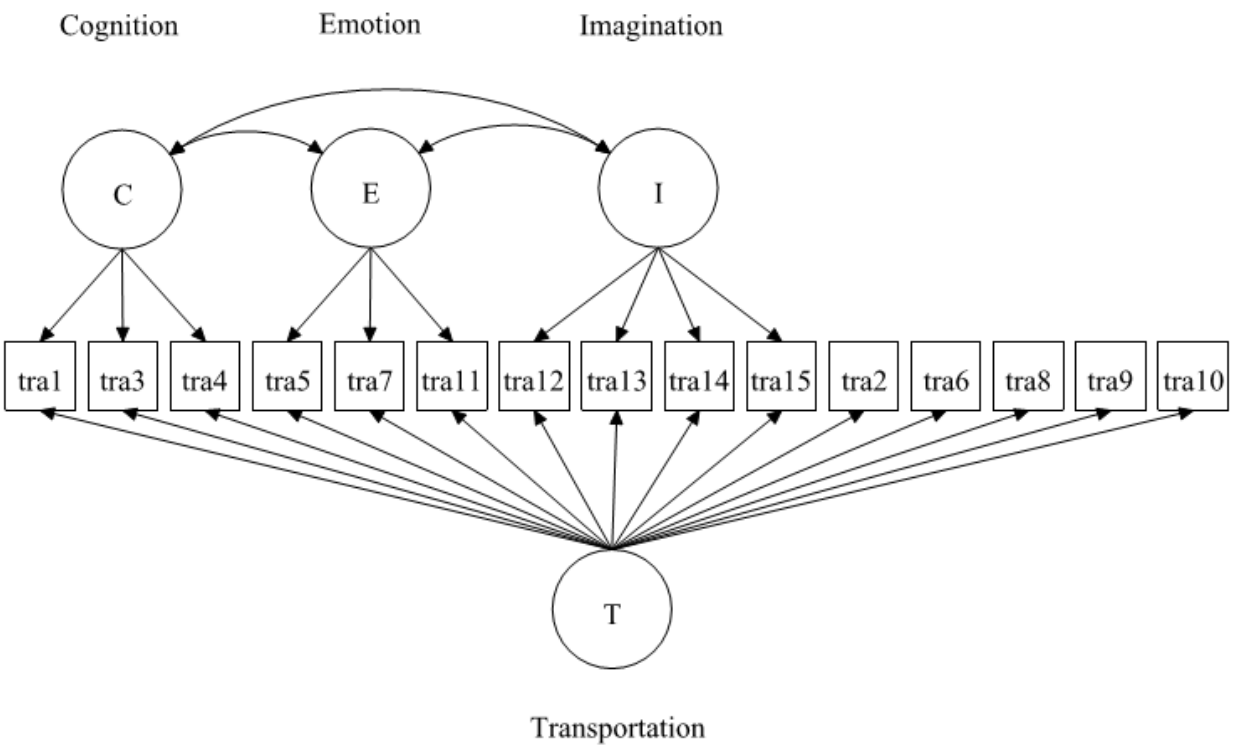

Figure 1. Bifactor specification for the facet structure of the original Transportation Scale

Table 1

Bifactor Loadings for the Original Transportation Scale (Study 1)

Item	Transportation	Cognition	Emotion	Imagination	h^2
1.	**.80 / .64**	**.33 / .48**	-.15 / -.14	.08 / .07	.77 / .66
2.	**.27 / .54**	-.54 / -.36	-.08 / .04	-.04 / -.15	.40 / .48
3.	**.75 / .64**	**.47 / .47**	.08 / .07	-.14 / .11	.79 / .71
4.	**.77 / .65**	**.37 / .73**	.07 / .03	-.01 / -.03	.75 / .95
5.	**.21 / .61**	-.36 / -.10	**.64 / .60**	-.05 / -.17	.49 / .71
6.	**.57 / .59**	.18 / .40	.13 / .00	.20 / .00	.46 / .51
7.	**.50 / .66**	.25 / .10	**.59 / .56**	-.06 / .04	.68 / .81
8.	**.25 / .39**	.31 / .10	.12 / .34	.43 / .22	.47 / .40
9.	**.87 / .73**	-.40 / -.03	-.02 / -.07	-.18 / -.10	1.00 / .55
10.	**.11 / -.01**	.16 / .11	.69 / .46	.20 / .06	.69 / .26
11.	**.14 / .28**	.02 / .01	**.75 / .51**	.04 / .18	.60 / .43
12.	**.72 / .67**	.36 / -.04	-.15 / -.14	**.28 / .72**	.75 / .90
13.	**.68 / .69**	.00 / -.13	.02 / .03	**.57 / .61**	.79 / .80
14.	**.49 / .54**	.03 / .10	-.10 / -.01	**.56 / .60**	.54 / .72
15.	**.09 / .18**	-.33 / .03	-.03 / .05	**1.00 / .73**	.98 / .61
	Variances explained as ω (Zinbarg et al., 2005)				
	72% / 73%	9% / 4%	1% / 6%	12% / 10%	
	Factor correlations				
	Emotional	.19 / .17			
	Imaginary	.25 / .45	.28 / .30		

Note. N = 68 / 111. Limited information factor analysis. Values on the left for the story *Christmas* and on the right for the story *Murder*. Hypothesized loadings are in bold. General factor is uncorrelated with facets. Item numbers correspond to Table 1 in Green and Brock (2000).

Table 2

Bifactor Loadings for the Transportation Scale - Short Form

	Study 1				Study 2a			
Item	Transportation	Cognition	Imagination	h^2	Transportation	Cognition	Imagination	h^2
3	**.61 / .74**	**.74 / .51**	.01 / .03	.92 / .81	**.59 / .73**	**.37 / .19**	.15 / .01	.55 / .57
4	**.75 / .89**	**.49 / .32**	-.01 / -.12	.80 / .89	**.64 / .79**	**.77 / .42**	-.01 / .00	1.00 / .80
6	**.98 / .80**	-.15 / -.10	-.08 / -.05	1.00 / .65	**.59 / .61**	.24 / -.01	-.09 / -.15	.40 / .40
7	**.57 / .63**	.23 / .15	.14 / .20	.41 / .48	**.94 / .72**	-.17 / .17	-.03 / .12	.92 / .57
12	**.63 / .70**	.18 / .02	**.46 / .50**	.70 / .74	**.43 / .93**	.06 / -.20	**.73 / .35**	.75 / 1.00
13	**.66 / .67**	-.15 / -.04	**.78 / .75**	1.00 / 1.00	**.53 / .74**	-.01 / .09	**.85 / .63**	1.00 / .99
	Variances explained as ω (Zinbarg et al., 2005)				Variances explained as ω (Zinbarg et al., 2005)			
	79% / 83%	8% / 3%	8% / 7%		71% / 87%	8% / 2%	13% / 4%	
	Factor correlation				Factor correlation			
	Imaginary	.33 / .26			Imaginary	.34 / .28		

Note. N = 68 / 111 for Study 1, N = 76 / 55 for Study 2a and N = 121 / 125 for Study 2b. Limited information factor analyses. Values on the left for the stories *Christmas* (Study 1), *Organ donation 1* (Study 2a) or *King* (Study 2b) and on the right for the stories *Murder* (Study 1), *Organ donation 2* (Study 2a) or *Chief* (Study 2b). Hypothesized loadings are in bold. Loadings marked with – were constrained to zero. Due to non-identification the cognition factor was not included in the models of Study 2b. General factor is uncorrelated with facets. Item numbers correspond to Table 1 in Green and Brock (2000).

Table 2 (continued)

Study 2b				
Item	Transportation	Cognition	Imagination	R^2
3	**.61 / .64**	**- / -**	- / -	.37 / .42
4	**.77 / .87**	**- / -**	- / -	.59 / .75
6	**.62 / .73**	- / -	- / -	.39 / .54
7	**.56 / .66**	- / -	- / -	.31 / .43
12	**.55 / .48**	- / -	**.65 / .57**	.73 / .56
13	**.52 / .65**	- / -	**.65 / .57**	.69 / .75
Variances explained as ω (Zinbarg et al., 2005)				
	74% / 81%	- / -	10% / 7%	

Table 3

Tests for Measurement Invariance of the TS-SF

Model	χ^2	*df*	*CFI*	*TLI*	χ^2/df	$\Delta\chi^2$	*df*	*p*	M
M1. Configural invariance	60.26	14	.96	.93	3.78				
M2. Factor loadings invariance	52.25	20	.97	.96	2.61	1.36	4	.85	M1
M3. Intercept invariance	108.42	49	.95	.97	2.21	56.51	33	.01	M1
M4. Partial intercept invariance	88.90	47	.96	.98	1.89	37.63	31	.19	M1
M5. Equal latent means	102.85	48	.95	.97	2.14	6.26	1	.01	M4

Note. *N* = 121 / 125. *CFI* = Comparative fit index, *TLI* = Tucker-Lewis Index, $\Delta\chi^2$ = Chi^2 difference, M = Comparison model. In model M4 the thresholds 5 and 6 of item 13 were allowed to vary freely in both groups.

Table 4

Descriptive Statistics for the Short and Long Transportation Scales (Study 2)

	M (*SD*)				Coefficient α (with 95% CI)		Cross-form correlation
	Long	Short	*d*	*Var* ratio	Long	Short	
Study 2a							
Text *Organ donation 1*	3.42 (0.84)	3.32 (1.15)	0.11	1.89*	.83 [.77, .87]	.84 [.77, .89]	.94*
Text *Organ donation 2*	3.61 (1.10)	3.60 (1.49)	0.01	1.83*	.88 [.82, .92]	.88 [.80, .93]	.94*
Study 2b							
Text *King*	3.87 (0.92)	4.22 (1.16)	-0.34*	1.60*	.83 [.78, .86]	.77 [.70, .82]	.89*
Text *Chief*	4.20 (1.02)	4.44 (1.34)	-0.20	1.70*	.84 [.79, .88]	.79 [.72, .84]	.95*

Note. CI = Confidence interval (based upon 10,000 bootstrap samples); *Var* ratio = Ratio of variances.

* $p < .05$

Table 5

Similarity Coefficients for the General Factor of the TS-SF

		Study 1		Study 2a		Study 2b
		Story 1	Story 2	Story 1	Story 2	Story 1
Study 1	Story 1					
	Story 2	.989				
Study 2a	Story 1	.940	.956			
	Story 2	.960	.984	.949		
Study 2b	Story 1	.982	.999	.963	.982	
	Story 2	.982	.993	.970	.967	.995

Appendix

Transportation Scale – Short Form (TS-SF)

Item No.	Item No. TS	Facet	TS-SF English	TS-SF German
1.	3.	Cognitive	I could picture myself in the scene of the events described in the narrative.	Ich konnte mich selbst in der Szenerie sehen, die in der Geschichte beschrieben wird.
2.	4.	Cognitive	I was mentally involved in the narrative while reading it.	Während des Lesens fühlte ich mich gedanklich in die Geschichte hineingezogen.
3.	6.	General	I wanted to learn how the narrative ended.	Ich wollte wissen, wie die Geschichte ausgeht.
4.	7.	Emotional	The narrative affected me emotionally.	Die Geschichte hat mich emotional berührt.
5.	12.	Imaginative	While reading the narrative I had a vivid image of *Katie*.	Während ich die Geschichte las, konnte ich mir *Katie* lebhaft vorstellen.
6.	13.	Imaginative	While reading the narrative I had a vivid image of *Joan*.	Während ich die Geschichte las, konnte ich mir *Joan* lebhaft vorstellen.

Note. Items were presented with seven-point response scales from 1 (*not at all*) to 7 (*very much*). Item numbers TS correspond to those provided by Green and Brock (2000, Table 1). Italicized names in items 12 and 13 need to be changed to the names of the main characters of the narrative.