

ARTICLE



<https://doi.org/10.1057/s41599-025-06341-2>

OPEN

I, ChatGPT: linguistic properties and human experiences of human- versus AI-generated stories

Markus Appel^{1✉}, W. P. Malecki², Tanja V. Messingschlager¹ & Julia R. Winkler¹

Generative AI is increasingly used to create texts, including fictional stories. Do stories generated by AI differ from stories created by humans regarding linguistic properties and recipients' experiences? To answer these questions, we first asked ChatGPT and 100 non-professional human authors (i.e., students) to create stories based on similar prompts (Study 1). Linguistic analyses showed that ChatGPT stories included fewer personal pronouns and fewer descriptions of relativity than human stories, but more positive emotions. In Study 2 ($N = 380$), naïve participants were randomly assigned to read an AI-generated or a human story from the pool of 200 stories from Study 1. No differences in novelty or entertainment experiences were found, but the readers of AI stories were less transported into the story world. Mediation analyses show that this difference can be attributed to ChatGPT's use of fewer personal pronouns. Differences in the use of literary techniques between AI and humans are discussed.

¹University of Würzburg, Würzburg, Germany. ²University of Wrocław, Wrocław, Poland. ✉email: markus.appel@uni-wuerzburg.de

Introduction

Ever since ChatGPT was made available to the general public in late November 2022, users worldwide have increasingly used AI chatbots based on large language models (LLMs), such as ChatGPT, to produce all kinds of texts. This includes fictional stories, a development that has been perceived as a challenge to the future of screenwriters and other entertainment professionals. Stories are omnipresent in people's lives and the widespread generation of stories by AI could have intriguing consequences for the experience of stories, as well as downstream effects on attitude change, the development of social-cognitive skills and other variables (cf. Green and Appel, 2024; Mar, 2018). Extant research (e.g., Messingschlager and Appel, 2024) suggests that AI authorship information—stories introduced as generated by AI (versus introduced as created by a human)—reduces the likelihood that recipients get deeply transported into a story world (*narrative transportation*, Gerrig, 1993; Green and Brock, 2000; Green and Appel, 2024). Little is known, however, about the potential that stories actually created by ChatGPT (or other AI-based programs) have on recipient experiences. To address this lacuna, two studies were conducted. In Study 1, we first developed 100 prompts (task descriptions) that included the instruction to write an entertaining short story. These 100 prompts served as instructions for 100 students who wrote one short story each, and for ChatGPT using the same prompts. An automated analysis of linguistic properties of the resulting texts was conducted (Linguistic Inquiry and Word Count, LIWC), with a focus on text characteristics we hypothesized to predict recipients' narrative transportation. In a subsequent experiment (Study 2), participants were randomly assigned to read one of the 100 human-created or one of the 100 AI-generated stories. Recipient experience in terms of narrative transportation, perceived novelty, enjoyment, appreciation, and perceived expertise were assessed. We further examined whether participants could attribute the stories correctly to the author (human or AI). Prior ChatGPT use served as a moderating variable. In our final set of analyses, the focal LIWC scores of the texts served as mediating variables to explain observed differences in the experience of stories written by humans and AI.

On the experience of AI-generated stories

ChatGPT has gained huge popularity for its ability to perform generative tasks without being specifically trained for them. Its large language model (LLM) is able to generate texts about almost every topic in varying forms and styles. One of these forms is creative writing, including fictional stories (Taecharungroj, 2023). ChatGPT performs these tasks based on text prompts users enter via a chat bar. In this process the human contributes an original idea, while the AI acts upon it. AI-generated stories and entertainment more generally could have a substantial impact on processes of content creation in the entertainment industries.

How do individuals respond to stories written by generative AI platforms such as ChatGPT? Theory and research suggest that variations in the experience of stories can be attributed to (a) the story itself, (b) individual differences (e.g., personality, knowledge, prior exposure), (c) situational variables, including source and paratext, and (d) the interplay among these factors (e.g., Green and Appel, 2024; Groeben, 1981; Valkenburg and Peter, 2013). The available empirical studies on the experience of AI-generated stories mainly focused on paratextual source labelling effects, that is, participants were exposed to the same text(s) but were told in one condition that the text was generated by AI, and in a second condition that the text was generated by a human (Messingschlager and Appel, 2024). Theoretical perspectives suggest a tendency for people to attribute superior creative

abilities to humans compared to AI (anthropocentrism, e.g., Millet et al., 2023; Messingschlager and Appel, 2025), and the concept of artistic creativity is strongly associated with being human (Chamberlain et al., 2018). Lower expectations for AI-created stories could translate to actual differences in the experience of stories due to AI source labelling (e.g., Tiede and Appel, 2020). Messingschlager and Appel (2024) focused on recipients' transportation into narrative worlds, a holistic state of attention, story-cued imagery, and affect (Gerrig, 1993; Gerrig, 2023; Green and Brock, 2000; Green and Appel, 2024). They found that contemporary fiction stories with an AI authorship label elicited less narrative transportation than the same stories labelled as human-created, at least for stories set in the here and now (rather than science fiction stories for which the authors expected and found smaller AI story labelling effects).

A core question, and the core of our empirical endeavor, pertains to the actual stories. Are stories generated by AI more absorbing and entertaining than stories created by human beings? Do readers find them more creative and aesthetically valuable? In other words, is AI a superior storyteller to human beings or the other way around? There are compelling reasons to believe in either possibility. Consider, first, that storytelling is a creative endeavor. An important aspect of creativity is divergent thinking, or the ability to associate remotely related concepts (Weiss et al., 2021; Wu et al., 2020; Zhang et al., 2020). As AI makes use of very large numbers of remotely related concepts, it is reasonable to assume that its creativity (a key aspect to storytelling) will also be superior, and there is even some empirical evidence to support that (Koivisto and Grassini, 2023).

Still another reason why AI might outperform human beings at storytelling is related to the fact that, like any creative endeavor, storytelling is a cognitively complex activity that demands simultaneously performing a variety of mental operations attentively. As suggested by controlled-attention accounts of activity (Frith et al., 2021), creative activities may therefore be negatively affected by mental fatigue, distraction, and various other factors that impair executive functions and to which humans are more susceptible than AI. Consistent with this, the results of a recent study comparing the performance of humans and AI on creativity tasks suggested that "humans were overrepresented in producing common or low-quality responses," indicating that "the weakness in human creative thinking, compared to AI..., lies in executive functions" (Koivisto and Grassini, 2023, p. 13601).

On the other hand, humans may outperform AI at storytelling insofar as writing quality stories demands good understanding of human behavior and mental processes, something which people are still more capable at than AI (Serikov, 2022). Consider emotions. Understanding human emotions is necessary for crafting the language of the story in such a way that its emotional tone will be found appropriate by readers and structuring the plot in such a way that the story engages them emotionally (Gordon et al., 2018; Winkler et al., 2023). While AI has been used to perform sentiment analyses of texts and detecting dominant emotional patterns in large literary corpora, (Kusal et al., 2021; Reagan et al., 2016), it has yet to demonstrate the capacity to understand emotions with the sophistication available to human beings (Li et al., 2023). This consideration is also important, given that most narratives created and enjoyed by humans are either about human characters or about characters who are anthropomorphic (Albee, 2015; Mar, 2018). The more their behavior and mental processes are made consistent with typical emotional patterns in humans, something at which people can be expected to outperform AI, the more believable the characters are, and the more absorbing and appreciated the story is going to be (El-Nasr et al., 2009; Saillenfest and Dessalles, 2014; Shirvani, 2019). More specifically, when prompted to produce an

entertaining story, AI-created stories could lack the poignant, bittersweet moments in life that elicit eudaimonic entertainment experiences (Oliver et al., 2021; Oliver and Bartsch, 2010).

In light of the above and related considerations, establishing whether AI can outperform humans at storytelling can only be established empirically.

Linguistic properties and the experience of stories

Differences between stories generated by AI and those generated by humans in terms of narrative transportation and related constructs could be due to systematic differences in the linguistic properties of the stories.

The first class of linguistic elements expected to predict the experience of stories are personal pronouns. Story writers can familiarize the reader with the characters and occurring events by telling the story from one or more perspectives within the narrative worlds. Personal pronouns, typically referring to the protagonists are an established way to clarify the concrete perspective taken (Who is telling the story? From which character's perspective is the story told?). Theory and initial evidence on the sentence comprehension level (e.g., Brunye et al., 2009; 2011; Hartung et al., 2016; Sanford and Emmott, 2012) suggest that the use of personal pronouns contributes to narrative transportation. Thus, we expected that stories with more personal pronouns would elicit higher narrative transportation and related experiences.

Second, emotional responses have long been recognized as a key part of the narrative experience (e.g., Oatley, 1999; Mar et al., 2011), and shifts between positive and negative emotional experiences tend to intensify narrative transportation (Nabi and Green, 2015; Winkler et al., 2023). We assume that the emotions experienced and expressed by story characters likely contribute to intense narrative experiences (e.g., Appel and Richter, 2010); thus, the emotionality in the text (positive and negative) should predict the experience of narrative transportation and related experiences.

Finally, we expect linguistic elements belonging to the categories of relativity and perceptual processes (Meier et al., 2018; Pennebaker et al., 2001) to facilitate transportation. Narratives are anchored in time and space (e.g., Dahlstrom, 2014), and transportation relies on readers' construction of a mental model of the story world and the causally connected narrative events (e.g., Busselle and Bilandzic, 2008). Vivid imagery of the story world and perspective taking with the characters are core facets of narrative transportation (Green and Brock, 2000; Green and Appel, 2024). Relativity words (prepositions and words that help to indicate space, time, and motions) may aid readers in the process of constructing a mental model of the narrative world and the chronological sequence of events. Further, verbalization of perceptual processes like seeing, hearing, or feeling may facilitate perspective taking, as they let readers witness how the character experiences the narrative world around them.

Study overview and predictions

To examine differences in the experience of AI-generated versus human-created stories, we first developed a viable sample of texts in which the goal of text production was identical for the AI and the human authors. To this end, 100 German students were asked to write an entertaining story, and the protagonist/topic of the story differed for each student. These 100 prompts were also used to generate German-language stories with ChatGPT. In a first set of analyses, we compared the resulting texts. We were specifically interested in textual differences on the dimensions we expected to influence narrative transportation and related reader experiences. Guided by theory and the LIWC linguistic text analysis dimensions (Meier et al., 2018), our pre-registered hypotheses were that

narrative transportation and related experiences were positively associated with positive emotion, negative emotion, personal pronouns (higher level category consisting of 1st, 2nd, 3rd person pronouns), perceptual processes (higher level category consisting of seeing, hearing, and feeling), and relativity (higher level category consisting of motion, space, time). In Study 1, we compared stories created by ChatGPT to stories created by humans regarding these linguistic dimensions.

In Study 2, we asked a large sample of recipients to read the stories created by humans or AI and to report on their experiential state during reading. Given the diverging lines of theory outlined above, we proposed several undirected mean differences hypotheses. We further assumed that the underlying craftsmanship of our student volunteers differed substantially, leading to a larger variance in recipient experiences of student volunteer stories as compared to AI-generated stories.

We expected a mean difference between stories created by AI versus stories created by human authors in terms of narrative transportation (H1), and higher variance in narrative transportation in the human authors condition as compared to the AI condition (H2). We further asked participants about the perceived novelty of the text, a key component of perceived creativity (along with usefulness, e.g., Runco and Jaeger, 2012). We expected a mean difference in perceived novelty between stories created by AI versus stories created by human authors (H3) and higher perceived novelty variance in the human-created stories (H4).

Much of the literature on the experience of narratives and entertainment media has focused on *enjoyment* and *appreciation* as two dimensions of experience (e.g., Oliver et al., 2021; Oliver and Bartsch, 2010). Enjoyment reflects the hedonic component of the entertainment experience whereas appreciation reflects the eudaimonic component of the entertainment experience (Oliver and Raney, 2011; Oliver et al., 2021). We expected a mean difference in enjoyment (H5), and higher variance in perceived enjoyment when reading the human-created stories (H6). As a directed hypothesis we expected that stories created by human authors elicit more appreciation than stories written by AI (H7), and we again assumed that the variance for this variable was higher when reading human-created stories (H8).

In addition to our hypotheses, we addressed several research questions: First, we examined the moderating role of recipients' prior experience with LLMs: Do the focal mean differences vary with recipients' prior experience with LLMs such as Chat GPT? (RQ1). We were further interested whether or not the AI-generated and the human stories were equally attributed to a professional author. Thus, our second research question was: Does story condition influence recipients' assessment of the author's expertise? (RQ2). Third, the correct attribution of the source was of interest: Can participants distinguish between AI stories and human stories? (RQ3). Fourth, we wished to gain insight into the association between source attribution and perceived expertise: Is recipients' assessment whether the story was written by AI or by a human related to perceived expertise? (RQ 4).

In a final step we analyzed whether potential differences in narrative transportation and related experiences between texts written by human authors and ChatGPT are mediated by the linguistic text factors outlined above. In other words, if differences in the experience of stories exist, could these be explained by differences in the linguistic properties?

Study 1: Generation and linguistic analysis of human and AI-written stories

Method. Our goal at the initial stage was to generate a sample of stories written by AI and stories written by humans. In terms of

sample size, prior research that examined the same stories, and manipulated authorship (the same stories were introduced as generated by AI versus introduced as created by a human) yielded differences in the magnitude of $d = 0.45$ and $d = 0.44$. (Messingschlager and Appel, 2024). Using these results as a starting point, an a priori sample size analysis (G^* power) to detect a mean difference of $d = 0.40$ between two groups, with $\alpha = 0.05$ and $\text{power} = 0.80$, amounted to 200 units of analysis (in our case: texts). Thus, we used 100 prompts that were intelligible to human authors and to ChatGPT, yielding 100 stories created by AI and 100 stories created by humans. The resulting 200 stories were analyzed using LIWC, with a theory-guided focus on linguistic properties that we expected to influence narrative transportation.

Participants. We initially recruited a sample of 100 undergraduate students who participated for course credit. The students were enrolled in a program that connects psychology, communication science, and computer science at a German university. Due to the fact that 10 of the participants' texts diverged substantially from the given task (see more in the "Procedures" section), an additional sample of 10 undergraduates were recruited in a second step. The final author pool consisted of 72 women, 37 men, and one person of non-binary gender who were between 19 and 30 years old ($M = 21.90$, $SD = 2.07$).

Software. The AI stories were generated between September 1 and September 13, 2023, using ChatGPT (based on GPT-3.5).

Procedure and prompts. The students were recruited to participate for course credit in a study on short story production. The study took place in a lab with groups of 2–8 students per session. Each student was placed in front of a computer. An MS Word document was provided that included the prompt. Each participant received a slightly different prompt. Note that prompts differed in their topics [example placed in parentheses below], all other parts of the instructions were identical for all participants. The wording was:

Please write a story that is as entertaining as possible. The story is about [a creative saleswoman]. The story must not be longer than 400 words. The story also needs a title. You have 50 min to complete this task. Please use the entire time to complete the task.

Each participant created one story. The computers were disconnected from the internet and the experimenter monitored that participants did not use ChatGPT or other AI with their smartphones. After the story was completed, the participants answered a brief demographics questionnaire and were dismissed. A total of 10 participants' texts diverged substantially from the specified task. Nine of the texts lacked a heading and one text was much longer than instructed. As a result, an additional 10 participants were invited to contribute short stories based on the prompts that had resulted in inadequate texts in the first round. All texts produced in this second round were adequate.

ChatGPT received the same 100 prompts our student sample had worked on. The wording was identical, except that the last two sentences in which the timeframe was explained were suspended, that is:

Please write a story that is as entertaining as possible. The story is about [a creative saleswoman]. The story must not be longer than 400 words. The story also needs a title.

All texts generated by ChatGPT matched our instructions. A list of the 100 topics can be found online in the OSF project (<https://osf.io/su3j6/>).

Linguistic analysis. We used LIWC and the DE-LIWC2015 dictionary (Meier et al., 2018) to calculate linguistic properties of each text. LIWC analyzes each word in the text and scores whether it belongs to a category specified in the dictionary. For each dictionary category, it calculates the percentage of words in the text that falls into the category. Given that the counts (words per category) and total text length constitute the LIWC scores, it is taken into account that longer texts have a higher likelihood to include words of a given category. Among the large set of variables available in the DE-LIWC2015 dictionary, we were particularly interested in the linguistic categories that we expected to be related to recipient transportation, that is, *positive emotion*, *negative emotion*, *personal pronouns* (higher level category, consisting of 1st, 2nd, and 3rd person pronouns), *perceptual processes* (higher level category consisting of *see*, *hear*, *feel*, e.g., "sehe [see]", "Klang [sound]", or "glatt [smooth]"), and *relativity* (higher level category consisting of *motion*, *space*, *time*, e.g., "Ankunft" [arrival], "unter" [below], "bisher" [hitherto]). The analysis of these categories (and no other category) was pre-registered (https://aspredicted.org/W3G_PJ5). The linguistic analyses were pre-registered before LIWC data generation but after Study 2 data collection.

Results and discussion. For the quantitative, inference-statistical comparison of the linguistic characteristics of the two sources (ChatGPT vs. human authors), we conducted Welch-tests, which do not require equal variances (and have been described to be preferable to Student t -tests, Delacret et al., 2017). Skewness and kurtosis of the main dependent variables were acceptable (Hair et al., 2018; see Supplement S1 for details) and no extreme outliers were observed. Alpha error probability was set to $p = 0.05$ and two-tailed tests were performed.

The texts were between 264 and 465 words long, with a mean of 362.14 words ($SD = 49.83$). On average, the texts included 14.12 words per sentence ($SD = 3.47$). A comparison between the texts written by humans and the texts written by AI revealed several linguistic differences. The texts written by humans were substantially longer ($M = 406.47$; $SD = 19.54$) than the texts written by AI ($M = 317.80$; $SD = 25.27$), $t_W(186.21) = 27.76$, $p < .001$, $d = 3.93$, 95% CI [3.45, 4.40]. Words per sentence did not differ (humans: $M = 14.11$; $SD = 1.70$; AI: $M = 14.13$; $SD = 3.47$), $t_W(144.11) = 0.05$, $p = .958$, $d = 0.01$, 95% CI [-0.27, 0.29].

Regarding our focal linguistic characteristics, ChatGPT-written stories included much more positive emotionality ($M = 5.46$; $SD = 3.60$) than stories written by humans ($M = 3.60$; $SD = 1.36$), $t_W(171.00) = -7.48$, $p < 0.001$, $d = -1.06$, 95% CI [-1.35, -0.76], whereas no difference regarding negative emotionality was observed (humans: $M = 1.87$; $SD = 2.08$; AI: $M = 2.08$; $SD = 1.45$), $t_W(170.65) = -1.18$, $p = 0.958$, $d = -0.17$, 95% CI [-0.44, 0.11]. Human stories included more personal pronouns ($M = 11.40$; $SD = 3.10$) than stories written by ChatGPT ($M = 8.93$; $SD = 2.14$), $t_W(176.04) = 6.56$, $p < 0.001$, $d = 0.93$, 95% CI [0.64, 1.22]. This difference was particularly remarkable for the 1st person pronoun (i.e., "I", "me" "my") which was used by humans much more often ($M = 2.81$; $SD = 3.96$) than by ChatGPT ($M = 0.21$; $SD = 0.44$), $t_W(101.47) = 6.54$, $p < 0.001$, $d = 0.93$, 95% CI [0.63, 1.22]. Humans used more 2nd person pronouns ($p < 0.001$; $d = 0.63$, 95% CI [0.35, 0.92]) whereas ChatGPT used slightly more 3rd person masculine/feminine pronouns ($p < 0.001$; $d = 0.29$, 95% CI [-0.57, -0.10]).

Moreover, descriptions of relativity were more prevalent in human stories ($M = 22.24$; $SD = 3.31$) than in stories by ChatGPT ($M = 20.13$; $SD = 3.19$), $t_W(197.73) = 4.97$, $p < 0.001$,

$d = 0.70$, 95% CI [0.42, 0.99]. We found no differences in the description of perceptual processes (humans: $M = 3.18$; $SD = 1.16$; AI: $M = 3.07$; $SD = 1.43$), $t_w(189.76) = 0.66$, $p = 0.509$, $d = 0.09$, 95% CI [-0.18, 0.37].

In sum, among the linguistic markers that we expected to elicit pronounced narrative transportation and related narrative experiences, personal pronouns and descriptions of relativity were more prevalent in human-created stories, whereas positive emotion words appeared less frequently in these narratives. Study 2 was conducted to examine whether the stories by our student volunteers or by Chat GPT yielded stronger narrative experiences and whether the textual differences would serve as explanatory variables.

Study 2: The experience of stories written by humans or by ChatGPT

Method. Study 2 followed an experimental design and made use of the stimuli created and analyzed in Study 1. The design and main effects analysis were pre-registered (https://aspredicted.org/3ZJ_7ST), as were the mediation hypotheses based on the linguistic text analysis reported in Study 1 (https://aspredicted.org/W3G_PJ5).

Participants. The number of participants was determined a priori. It was based on the sample size required to detect a mean difference of $d = 0.30$ between two groups, with $\alpha = 0.05$ and $\text{power} = 0.80$, amounting to 352 participants ($G^*\text{power}$). To account for the exclusion of careless responders, 406 participants participated via an invitation on Prolific. A total of 26 participants had to be excluded due to the following, pre-registered reasons: Seventeen participants completed the study in under 180 seconds, indicating low diligence, six participants did not summarize the study in meaningful German, indicating low diligence and/or low German skills, one participant failed the instructed response item (wording: “This is a control item. Please answer with 1 = *do not agree*.”), and two participants self-reported low diligence. Our final sample consisted of 380 participants (195 in the human author condition, 185 in the AI condition, see below) with an average age of 34.46 years ($SD = 11.82$, 166 female, 204 male, 10 non-binary or prefer not to say).

Stimulus material. The 200 stories generated in Study 1 served as our stimulus material. Participants were randomly assigned to read one story which was created by a human or ChatGPT (our main experimental factor).

Measures

Transportation: The participants reported the degree to which they were transported into the narrative world by answering the German version of the Transportation Scale-Short Form (TS-SF, Appel et al., 2015). It consisted of five items (e.g., *I was mentally involved in the narrative while reading it*; Cronbach’s $\alpha = 0.83$, $M = 4.78$, $SD = 1.21$). Note that instead of the two items that referred to the imagery of characters, we used one item assessing imagery in general terms (i.e., *I had a vivid mental image of the characters*). Unless indicated otherwise, all items went with a seven-point scale (1 = *not at all* to 7 = *very much*).

Novelty: Participants indicated the perceived novelty of the story with the help of four originality items by Moldovan and colleagues (2011). These items consisted of single adjectives (e.g., *original*, *unusual*; Cronbach’s $\alpha = 0.92$, $M = 4.03$, $SD = 1.53$).

Enjoyment and appreciation: We used three items each to measure enjoyment and appreciation based on Oliver and Bartsch

(2010). The items are frequently used in related research and yielded good psychometric properties (Schneider et al., 2019). The wording was adapted to fit the textual material (enjoyment, e.g., *It was fun for me to read this text*; Cronbach’s $\alpha = 0.92$, $M = 4.92$, $SD = 1.36$; appreciation, e.g., *The text was thought provoking*; Cronbach’s $\alpha = 0.87$, $M = 3.59$, $SD = 1.55$).

Perceived author expertise: Participants indicated the perceived author expertise with a single item (*The short story I just read was written by a professional writer*, $M = 3.31$, $SD = 1.54$).

Attributed authorship: human vs. AI: Likewise, a single item was used to assess whether participants attributed the text to an artificial intelligence ($M = 4.37$, $SD = 1.56$). The item was introduced with “Nowadays, short stories can be written by computer programs that use artificial intelligence (AI). What is your impression of the story you read in this respect?” and the item statement was “The short story I just read was written by a computer program (artificial intelligence)”.

Prior use of ChatGPT (or of other Large Language Models, LLMs): A single item was used to measure participants’ prior interactions with LLMs. It was introduced with “This question is about your personal experience with ChatGPT and other AI-powered chatbots for text generation. If you have no experience with such programs, please click on *Do not agree at all*” and was worded “During the course of the last year, I have used ChatGPT or other AI-powered chatbots intensively”, $M = 3.94$, $SD = 2.01$).

Procedure. When the study was advertised on Prolific and throughout the survey, no reference to artificial intelligence, ChatGPT or similar was made, until we asked for potential AI authorship at the end of the survey (see below). After giving informed consent, participants read one story, allocated randomly. Next, they indicated their level of transportation while reading, followed by enjoyment and appreciation, perceived novelty, perceived author expertise, and attributed AI authorship. Finally, we asked for socio-demographics and participants were debriefed. To detect careless responding, we included an instructed response item in the survey, assessed self-reported carelessness, and required participants to briefly summarize the study towards the end of the socio-demographics section.

Results and discussion. All requirements for conducting the quantitative analyses were met. Skewness and kurtosis of the main dependent variables were acceptable (Hair et al., 2018; see Supplement S2 for details) and no extreme outliers were observed. Alpha error probability was set to $p = 0.05$ and two-tailed tests were performed.

Relationships between dependent variables. We first inspected the zero-order correlations between the variables. As shown before (e.g., Johnson and Rosenbaum, 2015), transportation was closely and positively related to enjoyment and appreciation, as were both latter components of entertainment (see Table 1). Likewise, the more participants perceived the text to be novel, and to be created by an expert author, the higher the scores on transportation and both entertainment components. Interestingly, attributing the text to generative AI was associated with lower perceived expertise (RQ4), with lower transportation and lower enjoyment. These results are in line with theory and research that proposed and found a negative effect of AI authorship on transportation when authorship information was provided prior to reading (Messingschlager and Appel, 2024).

Table 1 Zero-order correlations between the dependent variables (Study 2).

	1	2	3	4	5
1 Transportation	-				
2 Novelty	0.370 (<0.001)	-			
3 Enjoyment	0.704 (<0.001)	0.418 (<0.001)	-		
4 Appreciation	0.690 (<0.001)	0.421 (<0.001)	0.575 (<0.001)	-	
5 Perceived Expertise	0.380 (<0.001)	0.323 (<0.001)	0.443 (<0.001)	0.284 (<0.001)	-
6 Attribution to AI	-0.178 (<0.001)	-0.062 (0.227)	-0.131 (0.011)	-0.071 (0.170)	-0.426 (<0.001)

N = 380. *P*-values are in brackets. Note that variables 1 through 5 were assessed before AI was ever mentioned.

The experience of stories written by humans and AI. Our hypotheses focused on recipients' experience of the stories. More specifically, we expected differences in means between stories written by human authors versus AI as well as differences in variance between both conditions. As the procedure and interpretation of mean differences depends on possible variance differences between the experimental groups, these analyses were conducted first. We expected higher variances in the human author condition than in the AI condition and tested these differences with Levene's tests. In contrast to what was expected in (H2, H4, H6, and H8), variances did not differ significantly for narrative transportation, $F(1, 378) = 0.76, p = 0.385$, originality, $F(1, 378) = 1.56, p = 0.212$, enjoyment, $F(1, 378) = 0.06, p = 0.802$, or appreciation, $F(1, 378) = 0.42, p = 0.515$. These results suggest that the experience of a story varies as much for the human-created stories as for the AI-created stories.

Our main questions pertained to mean differences in the experience of stories written by humans versus stories written by AI. We were further interested in moderation effects by participants' prior usage of LLMs such as ChatGPT (RQ1). Thus, our main effects analyses were accompanied by regressions with interactions between the factor author group (human = 0; AI = 1) and prior use of ChatGPT (*z*-standardized).

Our first analysis showed that narrative transportation was higher for human-authored texts ($M = 4.90, SD = 1.23$) than for AI-generated texts ($M = 4.65, SD = 1.18$), $t_W(378) = 2.04, p = 0.042, d = 0.21, 95\% CI [0.01, 0.41]$, in a simple regression $B = -0.25, SEB = 0.12, p = 0.043$ (H1, see Fig. 1). This difference was not moderated by prior ChatGPT use, $B = 0.14, SEB = 0.12, p = .252$.

In our next analysis, the same procedure was run for novelty evaluations as the dependent variable (H3). This time, no significant difference between human-authored texts ($M = 3.99, SD = 1.47$) and AI-generated texts ($M = 4.08, SD = 1.58$) was observed, $t_W(372.1) = -0.61, p = 0.545, d = -0.06, 95\% CI [-0.26, 0.14]$. No indication of a moderation by prior ChatGPT use was found, $B = 0.15, SEB = 0.16, p = 0.356$.

Similar results were obtained for enjoyment (H5): Human-authored texts ($M = 4.89, SD = 1.38$) yielded no significantly different enjoyment scores than AI-generated texts ($M = 4.94, SD = 1.35$), $t_W(377.7) = -0.35, p = 0.730, d = -0.04, 95\% CI [-0.24, 0.17]$. Prior ChatGPT use was not a significant moderator, $B = 0.18, SEB = 0.14, p = 0.214$. In our only directed main effects hypothesis (H7), we expected that human-authored texts evoked stronger appreciation than AI-generated texts. However, human-authored texts ($M = 3.59, SD = 1.57$) did not differ from AI-generated texts ($M = 3.59, SD = 1.53$), $t_W(377.8) = 0.01, p = 0.989, d = 0.00, 95\% CI [-0.20, 0.20]$. Again, prior use of ChatGPT did not serve as a moderator, $B = 0.27, SEB = 0.16, p = 0.089$.

In addition, we raised the question whether participants ascribed higher expertise to the human- or the AI-generated stories (RQ2). Human-authored texts ($M = 3.34, SD = 1.58$) did not differ from AI-generated texts ($M = 3.28, SD = 1.49$) in this

regard, $t_W(378) = 0.43, p = 0.667, d = 0.04, 95\% CI [-0.16, 0.25]$. However, we observed a moderation effect of ChatGPT experience, $B = 0.31, SEB = 0.16, p = 0.049$. Follow-up analyses showed that this interaction effect was based on different trends for participants low and high in ChatGPT experience that were not significantly different from zero at low or high ChatGPT experience scores ($M \pm 1 SD$): Participants with high ChatGPT experience scores ($M + 1 SD$) showed a non-significant tendency to ascribe higher expertise to the AI-generated text, $B = 0.23, SEB = 0.22, p = 0.291$, whereas the opposite, non-significant tendency was observed for participants with low ChatGPT experience scores ($M - 1 SD$), $B = -0.39, SEB = 0.22, p = 0.084$.

Attributions of the story to human authors or to AI. Next, we were interested whether participants could attribute the text they had read correctly to human or AI authorship (RQ 2). As the respective measure was continuous, the same statistical procedures as before could be applied. Variances between both story author groups did not differ, $F(1, 378) = 1.580, p = 0.210$. Although descriptively, AI-generated texts were more strongly ascribed to an AI source ($M = 4.52, SD = 1.60$) than human-authored texts ($M = 4.23, SD = 1.51$), they did not differ significantly overall, $t_W(373.90) = -1.84, p = 0.067, d = -0.19, 95\% CI [-0.39, 0.01]$. ChatGPT experience was not a significant moderator, $B = 0.29, SEB = 0.16, p = 0.072$ at the $p = 0.05$ -level. In partial support for the assumption that prior ChatGPT use facilitated the distinction between texts created by humans versus texts created by AI, follow-up analysis showed that participants with high prior ChatGPT use scores ($M + 1 SD$) attributed the AI-generated texts more strongly to AI than the human-generated texts, $B = 0.56, SEB = 0.23, p = 0.013$. For participants with low prior ChatGPT use scores ($M - 1 SD$), attributions to AI did not depend on whether the text was in fact generated by AI or by created by a human, $B = -0.01, SEB = 0.23, p = 0.951$. Johnson-Neyman statistics showed that on average, participants who scored 0.14 SDs above the mean or higher on prior ChatGPT use were more likely to correctly identify AI authorship.

Connecting linguistic test analyses results to recipients' narrative transportation. As outlined in Study 1, stories written by AI and stories written by humans differed regarding several linguistic variables. Could this be the reason underlying the result that human-created stories yielded more transportation than AI-generated stories? In other words, did the textual differences mediate the effect of authorship (AI vs. human) on narrative transportation?

The LIWC analyses had indicated that stories by humans included less positive emotionality than ChatGPT stories. Stories by humans further included more personal pronouns, and more indicators of relativity than stories by ChatGPT. Moreover, stories by ChatGPT were shorter, indicating that text length could be another plausible mediator explaining why human stories elicited higher transportation than ChatGPT stories.

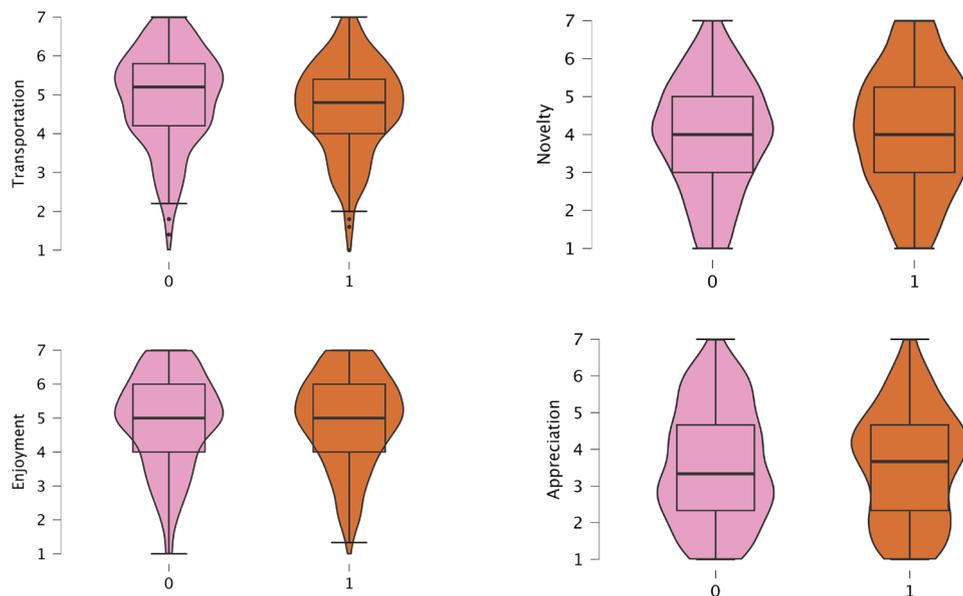


Fig. 1 Experience scores for stories written by humans and stories written by AI. Note. Distribution plots and boxplots. 0 = Human Author; 1 = AI Author. Significant differences were observed for Transportation; no significant differences were observed for the other dependent variables.

The following mediation analyses were conducted with PROCESS 4.2, model 4 (Hayes, 2022) with default specifications. Note that the causality structure in the model was in line with the experimental procedures, as linguistic indicators and experience were a consequent of story authorship and the experience measures were a consequent of the linguistic indicators (as part of the random story assignment).

Our first analysis pertained to text length. Theory (e.g., Gerrig, 1993; Green and Brock, 2002) suggests that for short stories or story fragments, longer texts should evoke more transportation (other story aspects being equal). However, to the best of our knowledge, no study so far had examined the influence of text length on narrative transportation. Zero-order correlations indicated that story word count affected transportation positively, $r(378) = 0.107$, $p = 0.037$. This weak positive association is an interesting result in and of itself. Text length, however, was not a significant mediator, effect estimate = -0.15 , $SE = 0.25$, 95% CI $[-0.64; 0.33]$.

Positive affectivity in the text was unrelated to recipients' transportation, $r(378) = -0.081$, $p = 0.116$, and the mediation analyses proper yielded no indirect effects (effect estimate = -0.05 , $SE = 0.07$, 95% CI $[-0.18; 0.08]$). Descriptions of relations showed small but significant associations with transportation, $r(378) = 0.112$, $p = 0.029$, but this linguistic variable was no significant mediator (effect estimate = -0.07 , $SE = 0.05$, 95% CI $[-0.17; 0.02]$).

A significant mediation effect was observed for personal pronouns (see Fig. 2). The complete bootstrapping results were as follows. Human stories contained more personal pronouns than ChatGPT stories, $B = -2.55$, $SE = 0.27$, $p < 0.001$. A higher number of personal pronouns in the text led to higher narrative transportation, $B = 0.07$, $SE = 0.02$, $p = 0.002$. The indirect effect was significant, effect estimate = -0.18 , $SE = 0.06$, 95% CI $[-0.31; -0.07]$. The total effect (IV- > DV) amounted to $B = -0.25$, $SE = 0.12$, $p = 0.043$. The direct (residual) effect was not significant, $B = -0.07$, $SE = 0.14$, $p = 0.618$. Note that these results remained virtually unchanged when the mediating variables were entered concurrently in one equation rather than separately.

General discussion

Summary and contribution. Generative AI is expected to change the workflows in creative industries and the cultural products

humans are exposed to (e.g., Anantrasirichai and Bull, 2022; Bohacek and Farid, 2024). Telling fictional stories is a particularly prominent part of the entertainment industries and fictional stories are important constituents of human cultures (Gottschall, 2012). The proliferation of generative AI has been accompanied by worries that it could worsen job prospects for human storytellers, as fictional stories are increasingly generated by AI (Appel et al., 2025). How do AI-generated stories compare to human stories? We evaluated stories told by AI and stories told by humans both in terms of linguistic properties and the ability to entertain and transport recipients into story worlds. Based on literary theory and research on story processing (Gerrig, 1993; Green and Appel, 2024) we gave AI (ChatGPT, GPT 3.5) and a sample of non-professional human storytellers the same task. We asked them to write an entertaining story based on the same 100 prompts. A linguistic analysis using LIWC (Meier et al., 2018) showed that ChatGPT stories included fewer personal pronouns, especially first-person pronouns, and fewer descriptions of relativity than human stories but more positive emotions. However, human and AI stories did not differ in terms of negative emotionality and words indicating perceptual processes. Importantly, our results further suggest that ChatGPT is not more proficient at accomplishing the storytelling task than our sample of students. On the contrary, students' stories were on average more transportive than AI stories. This difference was attributed to the more frequent use of personal pronouns by human storytellers.

This result is consistent with research that links the use of personal pronouns—particularly first-person pronouns—to perspective taking and mental imagery during reading (Hartung et al., 2016), and implies that the style of AI storytelling offers readers fewer opportunities to adopt characters' viewpoint or observe their thoughts and actions, which is crucial for the experience of transportation (e.g., Gerrig, 1993; Green and Brock, 2000). This interpretation is further strengthened by the potential relationships between pronoun use and narrative strategies, first-person and third-person voice in particular. First-person narratives, where the narrator is either the main protagonist or a minor character in the story, will generally be more conducive to the use of first-person singular pronouns than third-person narratives since at least one character in such a story will constantly report their actions and mental states in

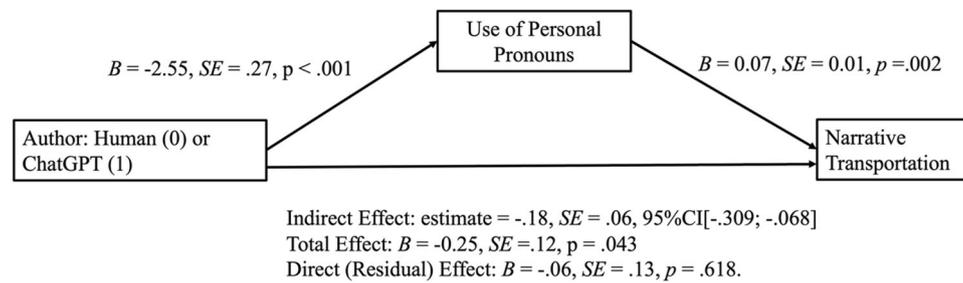


Fig. 2 Use of Personal Pronouns Mediates the Effect of Authorship on Narrative Transportation.

first-person (Bal, 2009). Some famous examples include Agatha Christie's *Who Killed Roger Ackroyd* ("It was just a few minutes after nine when I reached home once more. I opened the front door with my latch-key, and purposely delayed a few moments in the hall, hanging up my hat and the light overcoat that I had deemed a wise precaution against the chill of an early autumn morning", Christie, 1990, p. 163) and Margaret Atwood's *The Handmaid's Tale* ("Sometimes I listen outside closed doors, a thing I never would have done in the time before. I don't listen long, because I don't want to be caught doing it", Atwood, 1986, p. 10; cf. Garlick, 1992). In turn, third-person narratives with the so-called intrusive narrator, who expresses their opinion on the events in the story, very often directly addressing the reader, will be more conducive than third-person narratives employing a non-intrusive (or neutral narrator) to the use of both first- and second-person pronouns, both plural and singular. The intrusive narrator may adopt the position of both a singular subject and the so-called royal we, and address both a single reader and readers in general (Dawson, 2016). Also, both a lower frequency of first- and second-person pronouns and a greater frequency of third-person pronouns may be related to a preference for indirect speech and thought (*He thought he should do it/He said he should do it*) over direct speech and thought (*"I should do it," he thought/said*; e.g., Dancygier, 2019; Lucy, 1993; Vandelanotte, 2023). Of course, these are not the only possible factors underlying the difference in the use of personal pronouns between human and AI storytellers we observed, but they are worth further consideration in future research.

Consistent with extant theory and research that has identified transportation as an important mechanism of narrative effects, including entertainment experiences (Green and Appel, 2024), we found that transportation was positively associated with enjoyment and appreciation. However, this was not reflected in differences in enjoyment or appreciation between human and AI stories. This contrasts with recent research by Raffloer and Green (2025), who found that romance and science fiction stories written by AI—compared to stories written by graduate students—were enjoyed more (but results are less consistent for appreciation). Taken together, these findings suggest that the pleasure and meaning readers derive from AI-generated stories may to some extent be story- and genre-specific. On the other hand, transportation is affected by narrative qualities that transcend specific genres, such as artistic craftsmanship, verisimilitude, and narrative coherence (Green and Appel, 2024). Thus, using a large-scale stimulus set as in the present study, specific linguistic patterns of a LLM—such as a lower use of personal pronouns—might affect the experience of stories irrespective of genre. Although one could suspect that the relative inferiority of AI authors to tell transporting stories may diminish with the development of newer models, more recent studies comparing stories told by AI and humans yield conflicting results (cf. Chu and Liu, 2024; Raffloer and Green, 2025).

Regarding the identification of AI-generated stories, we found that on average, participants attributed AI-authorship to actual AI-generated stories to not significantly higher degree than to human stories. This is consistent with previous research demonstrating that individuals have difficulty correctly identifying AI-generated content, particularly text (Groh et al., 2024). Furthermore, we found that the more participants suspected AI to be the author of a story, the less transportation and enjoyment they experienced, and the lower they perceived the expertise of the author. These findings support the notion that people perceive AI as a less proficient storyteller than humans (Chu and Liu, 2024; Messingschlagler and Appel, 2024). Prior experience with AI was not linked to transportation or entertainment experiences.

Last, we found that variance in readers' experiences of stories by human writers is not significantly greater than variance between stories all generated by the same AI model. This was true for readers' transportation, originality, enjoyment, and appreciation. Interestingly, a recent study finds that stories that are written with the support of AI (but still with a human writer), are more similar to each other than stories written without any use of AI (Doshi and Hauser, 2024). Our stories by human writers might have had less variance due to the fact that we had a relatively homogeneous sample (all college students from the same major). Different levels of literacy and experience with writing narratives in human authors might lead to a more diverse set of narratives.

Our work contributes to and connects different fields of theory and research. First, our results add to the social scientific analysis of user responses to generative AI. We provide empirical evidence on the (yet limited) creative potential of generative AI in the field of storytelling (e.g., Appel et al., 2025; Epstein et al., 2023) and the linguistic properties associated with AI's lower capacity to engage recipients' minds (see more below). Second, our results contribute to theory and research on narrative processing (Busselle and Bilandzic, 2008; Green and Appel, 2024). The analysis of empirical user responses was based on a relatively large sample of different stories which enabled us to quantify linear associations between text properties and recipient transportation. Our results are relevant to theory and research on the antecedents of narrative transportation and may guide practitioners who wish to increase recipients' immersion into story worlds.

Limitations and directions for future research. Notwithstanding these contributions, our study has limitations. Our findings reflect the capabilities of one LLM, ChatGPT, based on GPT-3.5. Newer and more powerful LLMs have become available since our studies were conducted. With these developments, the differences between human stories and those generated by AI may diminish. At the same time, the human stories in our sample were written by non-professional authors (i.e., students), whose creative writing experience, abilities, and styles likely differ from those of

professional authors. Thus, future research should further investigate how AI stories compare to those written by professional human authors, both in terms of content (e.g., linguistic features) and how these stories are experienced. Future research may also investigate how the perception of AI versus human stories differs between non-expert and expert audiences (e.g., literary critics), because these audiences may use different evaluation criteria.

Moreover, we analyzed the stories regarding linguistic features that we expected to relate to the experience of transportation. However, the textual qualities of human and AI stories may differ in various other ways. For instance, future studies may investigate the presence of narrative arcs in human versus AI studies using the Narrative Arc feature of the LIWC-22 (see Boyd et al., 2020), a feature that was not available for German-language stories at the time of this study. Sentiment analyses can also reveal emotional arcs of stories (e.g., Reagan et al., 2016; Dale et al., 2023). Further, human and AI stories may differ in their use of literary techniques (e.g., foregrounding; van Peer et al., 2021). These features play a potential role in the experience and effects of narratives, and we hope to inspire further research on the ways in which human and AI storytelling differ and converge.

Conclusion

This research sheds light on the differences between stories authored by humans and AI by connecting their linguistic features to human experiences of narratives. By doing so, we contribute to an understanding of how storytelling and the experience of stories may transform in an age of generative AI. Our findings underscore that generative AI is able to produce art and cultural artifacts that are hard to distinguish from human generated art. However, we also provide an answer to the question of what stylistic features contribute to the relative inferiority of AI stories in terms of immersing their audiences, and therefore why AI stories at times may seem impersonal or bland.

Data availability

Materials, an online supplement, data, and analysis codes for the two reported studies can be found in the project's Open Science Framework repository, <https://osf.io/su3j6/>.

Received: 30 January 2025; Accepted: 18 November 2025;

Published online: 03 December 2025

References

- Albee A (2015) Our brains on stories—why stories work. In: Albee A (ed) *Digital relevance: developing marketing content and strategies that drive results*. Palgrave Macmillan US, pp 103–107. https://doi.org/10.1057/9781137452818_18
- Anantrasrichai N, Bull D (2022) Artificial intelligence in the creative industries: a review. *Artif Intell Rev* 55:589–656. <https://doi.org/10.1007/s10462-021-10039-7>
- Appel M, Gnamb T, Richter T, Green MC (2015) The transportation scale—short form (TS-SF). *Media Psychol* 18(2):243–266. <https://doi.org/10.1080/15213269.2014.987400>
- Appel M, Messingschlagler TV, Raffloer G, Reed P (2025) Generative artificial intelligence as creative artificial intelligence. In: Shackelford KE, Bowman ND (eds) *The Oxford handbook of media psychology*. Oxford University Press <https://doi.org/10.1093/oxfordhb/9780197689875.013.0016>
- Appel M, Richter T (2010) Transportation and need for affect in narrative persuasion: a mediated moderation model. *Media Psychol* 13(2):101–135. <https://doi.org/10.1080/15213261003799847>
- Atwood M (1986) *The Handmaid's Tale*. Harper Collins
- Bal M (2009) *Narratology: introduction to the theory of narrative*. University of Toronto Press
- Bohacek M, Farid H (2024) The making of an AI news anchor—and its implications. *Proc Natl Acad Sci USA* 121(1), e2315678121. <https://doi.org/10.1073/pnas.2315678121>
- Boyd RL, Blackburn KG, Pennebaker JW (2020) The narrative arc: revealing core narrative structures through text analysis. *Sci Adv* 6(32):eaba2196. <https://doi.org/10.1126/sciadv.aba2196>
- Brunyé TT, Ditman T, Mahoney CR, Augustyn JS, Taylor HA (2009) When you and I share perspectives: pronouns modulate perspective taking during narrative comprehension. *Psychol Sci* 20(1):27–32. <https://doi.org/10.1111/j.1467-9280.2008.02249.x>
- Brunyé TT, Ditman T, Mahoney CR, Taylor HA (2011) Better you than I: perspectives and emotion simulation during narrative comprehension. *J Cogn Psychol* 23(5):659–666. <https://doi.org/10.1080/20445911.2011.559160>
- Busselle R, Bilandzic H (2008) Fictionality and perceived realism in experiencing stories: a model of narrative comprehension and engagement. *Commun Theory* 18(2):255–280. <https://doi.org/10.1111/j.1468-2885.2008.00322.x>
- Chamberlain R, Mullin C, Scheerlinck B, Wagemans J (2018) Putting the art in artificial: aesthetic responses to computer-generated art. *Psychol Aesthet Creat Arts* 12(2):177–192. <https://doi.org/10.1037/aca0000136>
- Christie A (1990) *Agatha Christie: five complete murder mysteries*. Random House
- Chu H, Liu S (2024) Can AI tell good stories? Narrative transportation and persuasion with ChatGPT. *J Commun* 74(5):347–358. <https://doi.org/10.1093/joc/jqae029>
- Dahlstrom MF (2014) Using narratives and storytelling to communicate science with nonexpert audiences. *Proc Natl Acad Sci USA* 111(Suppl. 4):13614–13620. <https://doi.org/10.1073/pnas.1320645111>
- Dale KR, Fisher JT, Liao J, Grinberg E (2023) The shape of inspiration: exploring the emotional arcs and self-transcendent elicitors within inspirational movies. *Media Psychol* 26(6):767–789. <https://doi.org/10.1080/15213269.2023.2210841>
- Dancygier B (2019) Reported speech and viewpoint hierarchy. *Linguist Typology* 23(1):161–165. <https://doi.org/10.1515/lingty-2019-0004>
- Dawson P (2016) From digressions to intrusions: authorial commentary in the novel. *Stud Nov* 48(2):145–167. <https://doi.org/10.1353/sdn.2016.0025>
- Delacre M, Lakens D, Leys C (2017) Why psychologists should by default use Welch's t-test instead of Student's t-test. *Int Rev Soc Psychol* 30(1):92–101. <https://doi.org/10.5334/irsp.82>
- Doshi AR, Hauser OP (2024) Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Sci Adv* 10(28):eadn5290. <https://doi.org/10.1126/sciadv.adn5290>
- Epstein Z, Hertzmann A, the Investigators of Human Creativity (2023) Art and the science of generative AI. *Sci* 380(6650):1110–1111
- El-Nasr MS, Bishko L, Zammito V, Nixon M, Vasiliakos AV, Wei H (2009) Believable characters. In: Furht B (ed) *Handbook of multimedia for digital entertainment and arts*. Springer US, pp 497–528. https://doi.org/10.1007/978-0-387-89024-1_22
- Frith E, Kane MJ, Welhaf MS, Christensen AP, Silvia PJ, Beaty RE (2021) Keeping creativity under control: contributions of attention control and fluid intelligence to divergent thinking. *Creat Res J* 33(2):138–157. <https://doi.org/10.1080/10400419.2020.1855906>
- Garlick B (1992) *The Handmaid's Tale: narrative voice and the primacy of the tale*. In: Filmer K (ed) *Twentieth-century fantasists: essays on culture, society and belief in twentieth-century mythopoeic literature*. Palgrave Macmillan UK, pp 161–171. https://doi.org/10.1007/978-1-349-22126-4_13
- Gerrig RJ (1993) *Experiencing narrative worlds*. Routledge
- Gerrig RJ (2023) Processes and products of readers' journeys to narrative worlds. *Discourse Process* 60(4-5):226–243. <https://doi.org/10.1080/0163853X.2023.2177457>
- Gordon R, Ciocriari J, van Laer T (2018) Using EEG to examine the role of attention, working memory, emotion, and imagination in narrative transportation. *Eur J Mark* 52(1/2):92–117. <https://doi.org/10.1108/EJM-12-2016-0881>
- Gottschall J (2012) *The storytelling animal: How stories make us human*. Houghton Mifflin Harcourt
- Green MC, Appel M (2024) Narrative transportation: how stories shape how we see ourselves and the world. *Adv Exp Soc Psychol* 70:1–82. <https://doi.org/10.1016/bs.aesp.2024.03.002>
- Green MC, Brock TC (2000) The role of transportation in the persuasiveness of public narratives. *J Personal Soc Psychol* 79(5):701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Green MC, Brock TC (2002) In the mind's eye: transportation-imagery model of narrative persuasion. In: Green MC, Strange JJ, Brock TC (eds) *Narrative impact: social and cognitive foundations*. Lawrence Erlbaum, pp 315–342
- Groeben N (1981) The empirical study of literature and literary evaluation. *Poetics* 10(4-5):381–394. [https://doi.org/10.1016/0304-422X\(81\)90025-5](https://doi.org/10.1016/0304-422X(81)90025-5)
- Groh M, Sankaranarayanan A, Singh N, Kim DY, Lippman A, Picard R (2024) Human detection of political speech deepfakes across transcripts, audio, and video. *Nat Commun* 15:7629. <https://doi.org/10.1038/s41467-024-51998-z>
- Hair JF, Black WC, Babin BJ, Anderson RE (2018) *Multivariate data analysis*. Cengage

- Hartung F, Burke M, Hagoort P, Willems RM (2016) Taking perspective: personal pronouns affect experiential aspects of literary reading *PLoS ONE* 11(5):e0154732. <https://doi.org/10.1371/journal.pone.0154732>
- Hayes AF (2022) Introduction to mediation, moderation, and conditional process analysis: a regression-based approach (Vol. 3). Guilford Press
- Johnson BK, Rosenbaum JE (2015) Spoiler alert: consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation. *Commun Res* 42(8):1068–1088. <https://doi.org/10.1177/0093650214564051>
- Koivisto M, Grassini S (2023) Best humans still outperform artificial intelligence in a creative divergent thinking task. *Sci Rep* 13:13601. <https://doi.org/10.1038/s41598-023-40858-3>
- Kusal S, Patil S, Kotecha K, Aluvalu R, Varadarajan V (2021) AI based emotion detection for textual big data: techniques and contribution. *Big Data Cogn Comput* 5(3):3. <https://doi.org/10.3390/bdcc5030043>
- Li C, Wang J, Zhang Y, Zhu K, Wang X, Hou W, Lian J, Luo F, Yang Q, Xie X (2023) The good, the bad, and why: unveiling emotions in generative AI. *arXiv*. <https://doi.org/10.48550/arXiv.2312.11111>
- Lucy JA (1993) Reflexive language: reported speech and metapragmatics. Cambridge University Press
- Mar RA (2018) Evaluating whether stories can promote social cognition: introducing the Social Processes and Content Entrained by Narrative (SPaCEN) framework *Discourse Process* 55(5-6):454–479. <https://doi.org/10.1080/0163853X.2018.1448209>
- Mar RA, Oatley K, Djikic M, Mullin JB (2011) Emotion and narrative fiction: interactive influences before, during, and after reading. *Cogn Emot* 25(5):818–833. <https://doi.org/10.1080/02699931.2010.515151>
- Meier T, Boyd RL, Pennebaker JW, Mehl MR, Martin M, Wolf M, Horn AB (2018) “LIWC auf Deutsch”: the development, psychometrics, and introduction of DE-LIWC2015. Retrieved from <https://osf.io/tfqzcl/>
- Messingschlagler TV, Appel M (2024) Creative artificial intelligence and narrative transportation. *Psychol Aesthet Creat Arts* 18(5):848–857. <https://doi.org/10.1037/aca0000495>
- Messingschlagler TV, Appel M (2025) Mind ascribed to AI and the appreciation of AI-generated art. *N Media Soc* 27(3):1673–1692. <https://doi.org/10.1177/14614448231200248>
- Millet K, Buehler F, Du G, Kokkoris MD (2023) Defending humankind: anthropocentric bias in the appreciation of AI art. *Comput Hum Behav* 143: 107707. <https://doi.org/10.1016/j.chb.2023.107707>
- Moldovan S, Goldenberg J, Chattopadhyay A (2011) The different roles of product originality and usefulness in generating word-of-mouth. *Int J Res Mark* 28(2):109–119. <https://doi.org/10.1016/j.ijresmar.2010.11.003>
- Nabi RL, Green MC (2015) The role of a narrative’s emotional flow in promoting persuasive outcomes. *Media Psychol* 18(2):137–162. <https://doi.org/10.1080/15213269.2014.912585>
- Oatley K (1999) Why fiction may be twice as true as fact: fiction as cognitive and emotional simulation. *Rev Gen Psychol* 3(2):101–117. <https://doi.org/10.1037/1089-2680.3.2.101>
- Oliver MB, Bartsch A (2010) Appreciation as audience response: exploring entertainment gratifications beyond hedonism. *Hum Commun Res* 36(1):53–81. <https://doi.org/10.1111/j.1468-2958.2009.01368.x>
- Oliver MB, Raney AA (2011) Entertainment as pleasurable and meaningful: identifying hedonic and eudaimonic motivations for entertainment consumption. *J Commun* 61(5):984–1004. <https://doi.org/10.1111/j.1460-2466.2011.01585.x>
- Oliver MB, Raney AA, Bartsch A, Janicke-Bowles S, Appel M, Dale K (2021) Model of inspiring media. *J Media Psychol*. <https://doi.org/10.1027/1864-1105/a000305>
- Pennebaker JW, Francis ME, Booth RJ (2001) Linguistic inquiry and word count (LIWC): LIWC2001. Erlbaum
- Raffloer G, Green MC (2025) Of love & lasers: perceptions of narratives by AI versus human authors. *Comput Hum Behav Artif Hum* 5: 100168. <https://doi.org/10.1016/j.chbah.2025.100168>
- Reagan AJ, Mitchell L, Kiley D, Danforth CM, Dodds PS (2016) The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Sci* 5(1):31. <https://doi.org/10.1140/epjds/s13688-016-0093-1>
- Runco MA, Jaeger GJ (2012) The standard definition of creativity *Creat Res J* 24(1):92–96. <https://doi.org/10.1080/10400419.2012.650092>
- Saillenfest A, Dessalles J-L (2014) Can believable characters act unexpectedly?. *Lit Linguist Comput* 29(4):606–620. <https://doi.org/10.1093/lc/fqu042>
- Sanford AJ, Emmott C (2012) *Mind, brain and narrative*. Cambridge University Press
- Schneider FM, Bartsch A, Oliver MB (2019) Factorial validity and measurement invariance of the appreciation, fun, and suspense scales across US-American and German samples. *J Media Psychol Theor Methods Appl* 31(3):149–156. <https://doi.org/10.1027/1864-1105/a000236>
- Serikov AE (2022) Analysis of human behavior as a condition for creative artificial storytelling. In: Bylieva D, Nordmann A (eds) *Technology, innovation and creativity in digital society*. Springer, pp 449–461. https://doi.org/10.1007/978-3-030-89708-6_38
- Shirvani A (2019) Towards more believable characters using personality and emotion. Proceedings of the AAAI conference on artificial intelligence and interactive digital entertainment. 15(1):1. <https://doi.org/10.1609/aiide.v15i1.5253>
- Taecharungroj V (2023) “What can ChatGPT do?” Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data Cogn Comput* 7(1):35. <https://doi.org/10.3390/bdcc7010035>
- Tiede KE, Appel M (2020) Reviews, expectations, and the experience of stories. *Media Psychol* 23(3):365–390. <https://doi.org/10.1080/15213269.2019.1602055>
- Valkenburg PM, Peter J (2013) The differential susceptibility to media effects model. *J Commun* 63(2):221–243. <https://doi.org/10.1111/jcom.12024>
- van Peer W, Sopčák P, Castiglione D, Fialho O, Jacobs AM, Hakemulder F (2021) Foregrounding. In: Kuiken D, Jacobs AM (eds) *Handbook of empirical literary studies*. De Gruyter, pp 145–176. <https://doi.org/10.1515/9783110645958-007>
- Vandelanotte L (2023) Constructions of speech and thought representation. *WIREs Cogn Sci* 14(2):e1637. <https://doi.org/10.1002/wcs.1637>
- Weiss S, Steger D, Kaur Y, Hildebrandt A, Schroeders U, Wilhelm O (2021) On the trail of creativity: dimensionality of divergent thinking and its relation with cognitive abilities, personality, and insight. *Eur J Personal* 35(3):291–314. <https://doi.org/10.1002/per.2288>
- Winkler JR, Appel M, Schmidt M-LCR, Richter T (2023) The experience of emotional shifts in narrative persuasion. *Media Psychol* 26(2):141–171. <https://doi.org/10.1080/15213269.2022.2103711>
- Wu C-L, Huang S-Y, Chen P-Z, Chen H-C (2020) A systematic review of creativity-related studies applying the remote associates test from 2000 to 2019. *Front Psychol* 11: 573432. <https://doi.org/10.3389/fpsyg.2020.573432>
- Zhang W, Sjoerds Z, Hommel B (2020) Metacognition of human creativity: the neurocognitive mechanisms of convergent and divergent thinking. *NeuroImage* 210: 116572. <https://doi.org/10.1016/j.neuroimage.2020.116572>

Author contributions

MA and WPM had the original idea. MA headed the studies and analyses and wrote a first draft of the manuscript, with input by WPM, TVM, and JRW. All authors reviewed and agreed upon the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Ethics approval

Based on the regulations for conducting research with adult human subjects in Germany (as indicated by the German Research Fund DFG and the German Psychological Society, DGPs, <https://www.dfg.de/en/basics-topics/basics-and-principles-of-funding/good-research-practice>), no formal IRB approval was required. An automatic exemption by the Human-Computer-Media Institute and the Ethical Review Board of the institute applied. Both reported studies were conducted in full accordance with the Declaration of Helsinki. The studies followed the ethical guidelines of the APA and the DGPs.

Informed consent by the participants

Participants in both studies were provided with comprehensive information regarding the study before the study started. This including its purpose, procedures, the intended use of data (i.e., analysis in aggregated form and publication in an academic journal without disclosure of any personal information), and their right to withdraw at any time without negative consequences. For Study 1, written informed consent was obtained in-person between July 3, 2023, and July 21, 2023. For Study 2, written informed consent was obtained between November 28, 2023, and November 29, 2023, through the online questionnaire platform. On the introductory page of the survey, participants were required to read the study information and indicate their agreement by clicking the “Agree” button before being able to proceed to the questionnaire. By submitting their responses, participants confirmed their informed consent to participate in the study.

Additional information

Correspondence and requests for materials should be addressed to Markus Appel.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025