



Effects of a robotic storytelling intervention integrating music and sound effects on prejudice toward mental illness

Sophia C. Steinhäusser*, Miriam Semineth, Birgit Lugin

Socially Interactive Agents, University of Würzburg, Würzburg, BY 97074, Germany.

***Correspondence to:** Sophia C. Steinhäusser. Socially Interactive Agents, University of Würzburg, Würzburg, BY 97074, Germany. E-mail: sophia.steinhaeuser@uni-wuerzburg.de

Received: March 11, 2026 **Accepted:** June 25, 2026 **Published:** June 26, 2026

Cite this article: Steinhäusser SC, Semineth M, Lugin B. Effects of a robotic storytelling intervention integrating music and sound effects on prejudice toward mental illness. *Empath Comput.* 2026;2:202605. <https://doi.org/10.70401/ec.2026.0022>

Abstract

Aims: Mental illnesses affect millions worldwide. Despite greater public awareness, significant stigma remains. Anti-stigmatization interventions can help reduce prejudice by promoting education and meaningful contact with people with mental illness. Technology-based interventions may further support these efforts by simulating such contact through approaches such as storytelling. Social robots may offer additional advantages for storytelling-based interventions, especially through the integration of music and sound effects.

Methods: In multimodal robotic storytelling, non-speech sounds remain largely overlooked, despite their importance in related media. To address this gap, we compared a robotic storyteller narrating a story about a person experiencing a panic attack using only voice and bodily expression with versions integrating sound effects, background music, or both.

Results: While the addition of sounds did not affect recipients' prejudice differently than storytelling without them, adding both sound effects and music increased transportation and improved associative empathy. In contrast, adding only music decreased associative empathy. Although mediation was not indicated, the results revealed transportation as a predictor for empathy and prejudice, while being only minimally influenced by sound integration itself.

Conclusion: The effects of sound integration in a robotic storytelling intervention on recipients' prejudice were mixed, recommending either the combination of both sound types or complete omission. Transportation was indicated as a key lever for increasing empathy and decreasing prejudice that warrants further investigation. Future work is needed to gain deeper insights into robotic storytelling as an intervention tool for reducing prejudice, including work on transportation as a modifiable factor and the integration of pre-post-measurements.

Keywords: Social robots, robotic storytelling, empathy, transportation, mental health, stigma

1. Introduction

Mental health is a serious issue in today's society. "A mental disorder is characterized by a clinically significant disturbance in an individual's cognition, emotional regulation, or behaviour"^[1]. When asked, people report having only limited intimate contact with individuals who have a mental disorder^[2], although statistics range from nearly every third^[3] to every seventh^[4] individual worldwide suffering from a mental disorder. This perceptual gap is explained by the large number of affected individuals hiding their illness not only from the general public^[2], but also from their friends and families^[4]. Studies show that the prevalence of mental health disorders is constantly increasing^[3,5], so "that almost the whole population will at some time have direct experience of such a disorder, either in themselves or in someone close"^[2,6]. There are many types of mental disorders, with the most prevalent ones being anxiety disorders, depressive disorders, and stress^[5,7]. While the direct outcomes of mental disorders, such as distress, impairments of functioning in important areas, and disabilities, already hinder affected individuals in pursuing their goals^[4,8,9], mental illness is also associated with a secondary burden: stigma^[8].

"Stigma is a social phenomenon"^[10] that is characterized by labeling, stereotyping by association with negative attributes, separation, status loss, and discrimination^[10,11]. In doing so, mental illness is often seen differently than physical illness^[4], including being taken less seriously^[5] and judged medically unwarranted^[12]. Although the effectiveness of medication indicates a solvable problem with brain chemistry, mental illness is often seen as the 'fault' of affected individuals themselves^[4]. This results in not only the attribution



of responsibility, but also perceived incompetence and unpredictable dangerousness^[9], prejudices that build the core of stigma^[13], which lead to discriminating behavior such as avoidance, withdrawal, and segregation^[9,13]. In consequence, affected individuals do not declare mental disorders publicly for fear of, for instance, not getting hired^[4]. “Indeed, many people describe stigma as being worse than the condition itself”^[14].

While the reduction of stigma related to mental health cannot directly reduce symptoms associated with a specific disorder, it can have a positive effect on affected individuals’ well-being^[15] by improving social standing and reducing social distance, thereby supporting their overall well-being. To achieve this, attitudes and beliefs, in which stereotyping and distancing are rooted, have to be changed^[11]. An approach to correcting myths about mental illness is improving knowledge, for instance by enhancing mental health literacy^[6], while attitudes, particularly perceived dangerousness, can be improved by increasing familiarity^[2]. Classical interventions aim to increase familiarity through contact^[16], often performed through individual interventions in which affected people tell their personal story^[9]. While contact reduces stigma through increased knowledge, its effects are driven primarily by reductions in anxiety and increased empathy and perspective-taking^[16]. Such contact can be simulated by new technological tools such as chatbots, which are particularly effective at reducing fear and social distance stigma when using a first-person perspective^[15]. Alike, communication with virtually embodied agents can approximate this contact, resulting in stigma reduction comparable to traditional video interventions^[17]. Furthermore, virtual reality (VR) technology offers great advantages for stigma reduction interventions. “By allowing users to ‘walk in another’s shoes’ in simulated environments, VR is thought to evoke empathy and understanding more effectively than traditional media”^[18]. This positive effect of technology-based contact simulation was reported within multiple studies across diverse stigmatized groups^[18–20], including people with mental disorders^[21–23]; for an overview see^[24,25].

A relatively new medium for technological persuasive interventions is the social robot. Social robots are physically embodied agents^[26] capable of verbal and non-verbal communication, enabling natural and interpersonal interaction with humans^[27] and supporting a socially interactive focus^[28]. While social robots are heavily researched in the persuasive context of education^[29–32] and their integration as health technology is also gaining ground^[33–35], including their use in mental health care in the role of, i.e., therapy assistants, social companions, peers, playmates, social mediators, or instructors^[36], the deployment of robots in interventions aiming to reduce stigma and biases is still in its infancy. Although studies show that social robots are capable of improving attitudes towards and provoking behavior changes within important societal issues such as environmental sustainability^[32,37,38] or personal issues such as healthy nutrition^[30,39], there are only a few attempts to utilize this technology to benefit stigmatized groups. One approach to directly promoting equality among individuals belonging to a stigmatized minority using social robots is the concept of *Fair Proxy Communication*^[40,41]. Defined as “fairness-enhancing communication with robotic proxies”^[41], this approach uses social robots as avatars in telepresent social interactions such as job interviews to represent individuals that may otherwise be vulnerable to bias-driven decisions based on their visual cues on, for instance, gender or cultural ethnicity^[40,41]. While this approach may mitigate the consequences of stigma and bias, it leaves the underlying problem itself largely unaddressed.

To the best of our knowledge, yet only one single study had been carried out that tried to reduce stigma using a social robot: in their study Korner *et al.*^[42] implemented an interactive presentation on stuttering that included educational information, storytelling, practical tips, and interactive components in a social robot. The authors reported a significant improvement in participants’ attitude towards stuttering individuals from pre- to post-measurement, showing that social robots can help reduce stigma^[42]. Given the great societal importance of mental health stigma outlined above, it is crucial to examine whether this effect translates to other stigmatized groups, especially those diagnosed with mental disorders. Storytelling robots, in particular, have been shown to be persuasive^[43] and capable of evoking empathy^[44,45], a key emotion in classical stigma-reduction interventions^[16], which already has been shown to be able to foster prosocial behavior in interaction with social agents such as robots (for an overview, see^[46]). Their physical embodiment may provide an advantage over storytelling technologies previously used in the classical stigma reduction interventions described above, such as video interventions^[47,48]. Moreover, compared to technological interventions delivered by, for example, chatbots^[15], social robots are capable of multimodal communication, including not only human-like modalities such as speech and body language, but also non-biomimetic modalities such as the integration of images^[49,50], colored lights^[51,52], and non-speech sounds, i.e., sound effects and background music^[52,53]. As music in particular is known to enhance the emotional perception in other storytelling media such as film^[54–56], while sound effects can increase recipients’ attention, ease understanding^[57], and help to create mental imagery^[58–60], these positive effects could further benefit stigma reduction through robotic storytelling.

Accordingly, we examined the potential of robotic storytelling to influence stigma and prejudice toward individuals with mental disorders, with a particular focus on the added value of sound effects and background music. To this end, we compared four versions of a robotic storytelling about a person with a mental disorder: (1) the story told by the social robot *Pepper* using speech and emotional body language; (2) with added sound effects; (3) with added background music; and (4) with both sound effects and background music integrated. Given that anxiety disorders are among the most prevalent mental disorders^[3], we focused on anxiety disorders, specifically obsessive-compulsive disorder (OCD), characterized by obsessional intrusive thoughts^[61], whose symptoms can be effectively portrayed in a short story depicting a brief time frame.

2. Related Work

In the following, theoretical background and related works on stigma and prejudice as well as interventions on their reduction across diverse media including social robots are presented.

2.1 Interventions reducing stigma

Stigma is a social process based on social judgment^[12] and linked value^[62]. When encountering a stranger, we immediately form an impression of who this person is. In doing so, we construct what Goffman^[63] terms a *virtual social identity*, that is, an assumed identity based on available social cues and category memberships. This constructed identity does not necessarily correspond to the person's *actual social identity*, which comprises their real attributes and characteristics. If an attribute becomes apparent that deviates negatively from what is expected for the social category in question, the individual may be redefined in light of this attribute. In extreme cases, this can result in a reduction of that person from a usual to a discounted one. This attribute constituting a discrepancy between the virtual and actual social identity is called a stigma^[63]. Hence, stigma is based on the labeling of distinguishing differences that are associated with negative attributes, resulting in a separation between “us” and “them”^[11], ranging up to “dehumanization, threat, aversion, and sometimes the depersonalization of others into stereotypic caricatures”^[62]. Goffman describes three types of such stigma based on the type of the distinguishing attribute: 1) “abominations of the body, the various physical deformities”, 2) “blemishes of individual character”, including, among others, mental disorders but also homosexuality or unemployment, and 3) tribal identities such as “race, nation, and religion”^[63]. According to this categorization, stigma against mental illnesses is a stigma against one's personality or character, which is perceived as weak^[63], even though “the judgment is in some essential way medically unwarranted”^[12]. This interpersonal or *public stigma* includes prejudice, stereotypes, and discrimination, i.e., negative behavior^[9,14,64]. *Prejudice* refers to negative attitudes towards outgroup members^[13,62], respectively members of a particular outgroup with a specific distinguishing attribute^[65], or, in the case of prejudice towards people with mental illness, people with one or more mental disorders^[66]. Etymologically, the word stems from the Latin term *praejudicium*, meaning “prior judgment”^[13], describing the judgment based on group membership rather than personal attributes^[67]. Bizumin and Sheppard^[13] postulate four dimensions of prejudice towards people with mental illness, including 1) fear of affected people and thus a preference for social distance, 2) a perceived unpredictability of their behavior, 3) a perceived need for control over people with mental illness and restriction of their lives, and 4) malevolent beliefs that affected people are “physically, genetically, and psychologically inferior to people without mental illness” and should not receive help from societal resources^[68]. While in the past social scientists suggested prejudice being rooted in problematic personality traits, “social psychologists now consider stereotyping to be a normal (if undesirable) consequence of people's cognitive abilities and limitations, and of the social information and experiences to which they are exposed”^[62]. A stereotype is the cognitive element of prejudice^[13,69], describing rather generalized beliefs^[13,69] than attitudes. They “become prejudices when people develop negative emotions and evaluations” against stigmatized individuals^[9]. Finally, prejudice can result in discrimination^[9], an unjustified negative or harmful behavior toward an individual solely based on their membership in a stigmatized group^[13,69]. Taken together, stereotypes form the cognitive aspect of stigma^[13], prejudice describes its affective, attitudinal, and evaluative facets^[13,70], that drive behavior in the form of discrimination^[13]. Stigma includes prejudice “but extends to more general attributions about character and identity”^[62], although it is suggested that stigma and prejudice “describe a single animal”^[70].

Stigma against mental illness, particularly in terms of stereotypes of perceived dangerousness and discrimination through avoidance, is negatively correlated with familiarity, namely the knowledge about and experience with mental disorders^[2]. Providing information on mental illness and promoting contact with affected persons can challenge stereotypical expectations by disconfirming them^[71], for an overview see^[72]. Therefore, contact is a widespread and well-researched method for reducing prejudice via three routes: 1) increasing knowledge, 2) reducing anxiety, and 3) enhancing both empathy, hence, emotional responses^[73], and perspective taking, the attempt to see the world from another person's point of view^[74], with anxiety reduction and empathy enhancement being the strongest mediators^[16]. For instance, a field study examining contact with prisoners, in more detail murderers or convicted sex offenders with personality disorders, revealed an increase in empathy as well as decreased prejudice after five hours of contact^[75]. This stigma reduction can also be achieved by simulating contact through fictional stories such as in vignette studies^[15]. “Stories engage our minds and change the way we see the world”^[76]. In doing so, traditional storytelling, “characterized by a single teller addressing an audience using speech, physical gestures, possibly props, and non-speech sounds”^[77], has already been shown powerful within informal learning settings, such as with social stories for children fostering positive social behavior and decreasing aggressive behavior^[78]. In fighting prejudice, storytelling disconfirming stereotypes about African Americans in the form of a radio show was reported to improve perspective taking and decrease stereotypic perceptions^[72]. Moreover, animated storytelling videos were shown to decrease addiction stigma^[79] as well as transphobia^[80] in the short-term. Animations presented in VR were reported to successfully reduce stigma against people with anxiety and depressive disorder from pre- to post-intervention, with the effect sustaining to the one-week follow-up^[23]. Lastly, traditional oral storytelling was shown to effectively reduce stigma against people with a human immunodeficiency virus diagnosis directly after the intervention as well as six months later^[81]. These results reveal the great potential of storytelling as a substitute for direct contact with outgroup members through diverse media, reflecting findings on the persuasiveness of stories reported in general health communication^[82].

Next to empathy, transportation is an important variable in storytelling that positively influences changes in attitude. Transportation is defined as the recipients' absorption into the story, including an "integrative melding of attention, imagery, and feelings"^[76], and can occur across various media^[76]. Transportation can benefit persuasion and suppress counterarguments (for an overview see^[83]), leading to more story-consistent beliefs and less disbelief^[76]. With regard to story characters, transportation increases liking for them and subsequently reinforces beliefs that align with their experiences and perspectives as presented in the story^[76]. Finally, transportation has been shown to correlate positively with empathy, to promote prosocial behavior consistent with the narrative^[84], and to be negatively related to stigma^[23]. However, in the context of persuasive health messages, instrumental arguments can reduce transportation and increase psychological reactance. Thus, affective arguments are recommended^[85]. While multiple meta-analyses^[82,86] recommend the first-person perspective for written persuasive storytelling, Christy^[87] reported a beneficial effect of third- and second-person perspective compared to first-person perspective in written storytelling for improving attitudes towards gay people, although no differences in transportation or identification were obtained. For audiovisual media, the effect of perspective seems to be even less clear^[82]. Regarding the presentation medium, Zhuang and Guidry^[86] found no significant differences between written texts and videos or their combination in their stigma-reducing effects, suggesting the transferability of the approach between different media types^[86]. Nonetheless, adding pictures to text-based health messages can improve comprehension, recall, and adherence through increased attention^[88,89]. In stigma reduction, particularly emotion-conveying visual material may benefit intervention outcomes, given that the correct understanding of another person's distress cues, such as a fearful facial expression, enhances empathy and increases prosocial behavior^[90]. Also, other additional modalities could further benefit intervention effects. For example, music has been shown to influence prosocial behavior. Wu and Wang found that presenting sad music while participants read a story about a person in need increased their willingness to help compared to happy music or white noise, probably due to enhanced empathy^[91]. Within anti-stigma interventions, emotionally moving background music was reported to be a strong predictor for both empathy and reflective thoughts, although no significant effect on attitudes was found^[92]. In contrast, advantages of increased attention due to multimodality in stigma-reduction interventions have yet not been investigated. Taken together, storytelling across various media, whether written, animated, or traditionally orally delivered, holds substantial potential for anti-stigma interventions, with additional modalities such as music offering further opportunities to amplify their impact.

2.2 Non-speech sounds in storytelling media

The perception and interpretation of sound have played a vital role throughout human history, from early survival, when identifying sounds as potential prey or predators was essential, to modern contexts such as navigating traffic^[93,94]. Similarly, music is omnipresent in daily life, used to influence consumer behavior in retail settings and to support both physical and mental health treatments^[95]. Unsurprisingly, music and sound effects have long been intertwined with storytelling. Early oral storytelling incorporated musical elements^[96], which were later adapted for radio plays and films in the 1920s^[97,98]. Sound effects, originating in the theater^[60], were subsequently adapted in film and later in audiobooks in the 2000s^[97]. Nowadays, storytelling media soundtracks, such as those of films, "typically consist of music, speech, and sound effects"^[99].

Sound effects serve both orienting and descriptive functions, helping to create and sustain imagined environments^[60,94] by reconstructing reality through explicit sounds that represent, for example, objects or weather conditions^[100]. These sound effects can appear as short and specifically placed sounds, such as a barking dog, or as layered background ambience, such as echoing conversations and footsteps in a large city hall. Both approaches can also be mixed^[101]. Sound effects have been shown to improve immersion and focus on a story, making storytelling more enjoyable by strengthening mental imagery^[58,100]. When applied in persuasive communication, these effects can further enhance attitudes and memory performance^[59]. Beyond their merely diegetic functions, sound effects can also evoke emotional responses^[98,101], such as the tension created by the slow creak of an opening door^[101].

"Unlike speech and sound effects, music plays no part in the story as music per se"^[99], but functions as emotional guidance^[102]. Music, in general, has the high potential to express and evoke emotions^[103]. In doing so, it can influence human perception and information processing when combined with visual stimuli^[104,105]. For example, Jeong *et al.*^[106] found that happy faces presented alongside sad music were rated less happy, whereas the same faces paired with happy music were rated as happier, and vice versa. Alike, interpretations of social contexts were shown to be affected by music in terms of, *inter alia*, perceived valence and pleasantness^[104]. This cross-modal effect also extends to storytelling media: especially in film, emotionally charged background music has been shown to bias interpretation and memory in an emotionally congruent direction^[107-110], even affecting recipients' physiology^[95]. Part of this effect is explained by music's ability to evoke emotionally congruent thoughts, although the strength of this effect depends on familiarity with and liking of the musical piece^[111]. A widely accepted framework for understanding these phenomena is the Congruence-Association Model (CAM)^[112]. The CAM proposes that information congruent across modalities, such as the structural overlap of video and music (congruence), is processed with priority. This facilitates the aggregation of meanings from both channels (association), allowing "associations from the music [to] be ascribed to that focus of attention"^[112], ultimately resulting in an emotionally congruent interpretation of the combined audiovisual experience. The structural accordance further improves the perceived presence and transportation of recipients^[55,113], whereas visual storytelling without music was shown to decrease transportation into the story, causing less identification with the protagonist^[113]. Investigating the effects of sound effects, music, and

their combination, or omission of both, on movie reception, Kock and Louven^[114] found that music alone reduced the immersion range, while sound effects alone diminished suspense, compared to presenting both together. Based on these findings, they recommended using a combination of music and sound effects for audiovisual storytelling.

2.3 Multimodal robotic storytelling as an intervention tool

Social robots are a rather new but highly promising medium for storytelling. Being physically embodied^[26], these agents are designed for natural and interpersonal interaction with their human users, utilizing speech but also non-verbal cues such as body language or facial expressions^[27]. In contrast to service or industrial robots, social robots have a social interface^[115], so that they can support humans with a socially-interactive focus^[27,28], for instance, in educational contexts^[31,116] or aging^[117]. In doing so, social robots bear advantages over virtual agents due to their physical embodiment, offering richer interaction and sharing the user's interaction space^[118]. Based on their multimodal communicative abilities, robots can serve as storytellers, creating narrative experiences that are comparable to traditional audiobooks^[119]. Furthermore, both social robots in general^[46] and robotic storytellers in particular^[120] have been shown to be able to elicit empathy in their recipients. Initial intervention studies further suggest their potential effectiveness: for example, a robotic storytelling intervention against bullying improved children's empathy compared to a loudspeaker-based storytelling^[45]. To the best of our knowledge, however, robotic storytelling has not yet been examined in the context of mental illness stigma-reduction interventions. These early findings provide a promising foundation for future research exploring the potential of robotic storytellers in this domain.

Moreover, social robots offer not only human-like modalities such as voice and body language, but are also capable of integrating non-biomimetic modalities, including music^[121,122], which could further benefit anti-stigma interventions (see Section 2.1). Although research on additional sound integration in human-robot interaction is still in its infancy compared to, e.g., speech^[123], music has been shown to improve enjoyment of and transportation into robotic storytelling^[53], with the latter being an important key lever in stigma reduction (see Section 2.1). Moreover, music has been shown to facilitate emotion recognition from robotic expressions^[124]. As accurate recognition of negative emotions is a reliable predictor of prosocial behavior^[90], music accompanying a stigma-reducing storytelling intervention might amplify empathy and thus stigma changes. Given that robots can not only automatically integrate music into their storytelling^[125] without the need for additional devices (as compared to humans, see^[77]), but can also provide consistent telling over and over again (as compared to humans, see^[126]), they are a promising tool for providing standardized anti-stigma interventions and need further investigation.

3. Contribution

Facing the need for mental health awareness in today's society, we investigated the potential of multimodality integrating both human-like and non-biomimetic modalities in robotic storytelling to raise this awareness in recipients' and to improve their attitudes toward concerned persons. Therefore, we implemented a social robot telling a story of a person experiencing a panic attack accompanied by intrusive thoughts. To examine the efficiency of non-speech sounds in improving the influences on awareness and attitudes, we tested four variants of our storytelling: (1) using both sound effects and background music, (2) using only sound effects, (3) using only background music, or (4) omission of non-speech sounds. Both background music^[127] and sound effects^[100] can increase recipients' attention to storytelling media. Moreover, background music can have a positive effect on recipients' transportation into the story^[113] as well as the related concepts of immersion^[55] and cognitive absorption^[128]. However, as reported by Kock and Louven^[114], sound effects as well as the combination of background music and sound effects outperform the addition of music only. Therefore, we postulate the following hypotheses:

H1a: Attention is higher during the storytelling with non-speech sounds compared to that without.

H1b: Attention is higher during the storytelling including both sound effects and background music compared to those integrating only one type of non-speech sound.

H2a: Transportation is higher for storytelling with non-speech sounds compared to that without.

H2b: Transportation is higher for storytelling including both sound effects and background music compared to those integrating only one type of non-speech sound.

Both the increased transportation^[84] as well as the improved recipients' perception of the protagonist's emotions^[73], might increase their empathy for them. In addition, especially sound effects can increase attention during story reception^[57]. On this basis, we postulate the following hypothesis and research question:

H3: Empathy is higher during the storytelling with non-speech sounds compared to that without.

RQ1: Does the combination of sound effects and background music improve the positive effects of the individual non-speech sound types on empathy?

Combining modalities, such as audio and visuals, improves persuasive effects^[129], especially for narrative persuasive messages^[130].

Also for robotic storytellers, improved persuasiveness was obtained when adding emotionally matched facial expressions to the telling^[43]. Furthermore, multimodality was shown to improve sympathy and prosocial intentions while decreasing prejudice toward the target of hate speech^[131]. Although the addition of non-speech sounds has yet not been tested in terms of enhanced persuasion, based on the related works on multimodality in persuasive communication, we expect lower prejudice in recipients of the storytellings with added non-speech sounds.

H4a: Prejudice is lower after receiving the story with non-speech sounds compared to that without.

H4b: Prejudice is lower after receiving the storytelling including both sound effects and background music compared to storytelling integrating only one type of non-speech sounds.

Lastly, transportation was shown to improve the liking of story characters^[76] and empathy building^[84] as well as to reduce negative cognitive activities such as disbelief^[76], paving the way for the reduction of stigma. According to our focus on storytelling, we exploratorily analyzed whether transportation mediated our intervention's effects on both empathy and stigma reduction.

H5a: Transportation mediates the sound conditions' effect on empathy.

H5b: Transportation mediates the sound conditions' effect on prejudice.

4. Materials

A story about a person suffering from OCD was specifically written for this study. After pretesting the story in an online survey, it was implemented using the *Pepper* robot to realize four versions of robotic storytelling: (1) using both sound effects and background music, (2) using only sound effects, (3) using only background music, or (4) omission of non-speech sounds.

4.1 Story concept

We chose a female student as protagonist, named "Cora", for our story to facilitate identification with our predominantly young and female sample. The story (The whole story can be found at <https://go.uniwue.de/mentalhealth-material>) describes an everyday situation of a university student who suffers from OCD. The disorder manifests itself in obsessive thoughts that trigger states of anxiety and panic attacks. The protagonist has been in therapy for many years and is therefore aware of coping mechanisms that help her handle her condition.

The story was divided into seven paragraphs that describe Cora's journey to university on a warm summer's day. It contains 1,350 words, which translates to a duration of approximately ten minutes when told. The plot of the story is summarized as follows.

On her way to the bus stop, the protagonist accidentally collides with a child who runs off laughing. Shortly afterward, she is struck by an intrusive thought, "What if I was a pedophile?", which triggers intense anxiety despite her therapeutic knowledge that such thoughts are automatic and unfounded. The pedophilia-themed intrusive thought was selected because OCD-related intrusive thoughts frequently involve "deviant or non-normative sexual thoughts, as defined by Western cultural standards, such as the fear of developing an attraction to children"^[132]. In contrast to more commonly recognized OCD themes, such as contamination or checking behaviors, this theme was chosen to portray a particularly stigmatized and distressing form of obsessive thought^[133]. The protagonist's fear escalates when an ambulance passes by, leading her to worry that the child was hurt. While rushing to the bus, she feels a panic attack building, but once on board, after a boy asks whether she is okay, she applies coping strategies learned in therapy and gradually regains control. Arriving at the university, the protagonist is proud of her progress and her ability to manage her chronic condition. The story highlights typical thinking processes and emotional states that someone suffering from OCD might experience during an everyday situation. The intrusive obsessive thought was specifically chosen to be shocking to the recipients since these kinds of thoughts are often disturbing and violent in nature, causing high levels of anxiety and discomfort in affected parties^[134]. The story had a linear structure and offered no points of interaction for the recipient. The story was proofread by three persons. Moreover, to ensure the accuracy and appropriateness of the mental health content, the stimulus material was reviewed by a staff member of the Counseling and Information Service (KIS: <https://www.uni-wuerzburg.de/en/equity/kis/>) at the University of Würzburg.

4.2 Pretesting the story

In order to test the suitability of the story in terms of transportation and empathy elicitation, an online prestudy was carried out.

4.2.1 Measures

Transportation was assessed using the *Transportation Scale - Short Form* (TS-SF)^[135] comprising six items such as "I wanted to learn how the narrative ended" answered on a seven-point Likert-scale anchored by 1 - "not at all" and 7 - "very much". Appel *et al.*^[135] reported alpha reliability of .80 to .87, alpha for the current sample was .89.

Empathy was measured using the *State Empathy* scale by Shen^[136]. The scale consists of three dimensions comprising four items

each: (1) *Affective Empathy*, sharing the protagonist's feelings (e.g., "The character's emotions are genuine."), (2) *Cognitive Empathy*, understanding the protagonist's motivations and perceptions (e.g., "I see the character's point of view."), and (3) *Associative Empathy*, identification with the protagonist (e.g., "When watching the message, I was fully absorbed."). We adapted the third scale to better align with the context of robotic storytelling by replacing the term "the message" with "the story" and the term "the character" with the protagonist's name ("Cora"). All items were measured on a 5-point Likert scale anchored by 0 = "not at all" and 4 = "completely". Alpha reliabilities of the original scale reported by Shen^[136] were .83 to .91 for *Affective Empathy*, .86 to .91 for *Cognitive Empathy*, .82 to .92 for *Associative Empathy*, and .92 to .93 for the whole scale. For the current sample, alphas of .84 for *Affective Empathy*, .77 for *Cognitive Empathy*, .76 for *Associative Empathy*, and .91 for the whole scale were computed.

Personal relevance of the OCD topic was measured using the Personal Relevance Scale^[137]. The scale comprises four items such as "To what extent (if any) does the treated topic make you think of current situations or events in your own life?", which are answered on a five-point Likert scale anchored by 1 = "not at all" and 5 = "very much". Ellard *et al.*^[137] does not provide reliability measures, and as the scale was only used to exclude participants in our study but not for further analyses, we did not compute them either.

4.2.2 Procedure and exclusion criteria

Participants were acquired using the university's participant pool. In the study advertisement, a trigger warning concerning the topic of mental illness and OCD was provided. When entering the survey hosted via *LimeSurvey*, participants again received a trigger warning before they provided informed consent. Furthermore, we provided a list of contact points such as the crisis hotline or the university's internal psychotherapeutic counseling.

For ethical reasons, participants with current or past experiences with OCD, either in themselves or in a family member or friend, were excluded from the study to avoid triggering distress or exacerbating symptoms during the storytelling intervention. This exclusion was implemented through an aptitude screening conducted prior to the actual survey. Participants were asked whether they had ever received an OCD diagnosis (yes/no/prefer not to disclose), completed the personal relevance scale, indicated whether they met the minimum age requirement (yes/no), and confirmed their willingness to participate after reading the trigger warning (yes/no). Participants were excluded if they reported an OCD diagnosis, scored above the midpoint on the personal relevance scale, or chose not to disclose diagnostic information. In this case, they were informed about their exclusion and thanked for their willingness to participate. In doing so, individuals with high familiarity with OCD were excluded from the study. Data collected in the aptitude test were automatically analyzed during the participation and deleted prior to the main analyses.

Proceeding to the actual survey, participants first read the text version of the story described in [Section 4.1](#). Afterwards, they filled in the questionnaires on transportation and empathy, provided demographic information (age, gender), and were invited to provide comments. Last, the list of contact points was presented again.

The prestudy took about fifteen minutes to complete.

4.2.3 Participants

A total of 68 individuals participated in the pretest. Based on the screening results, 24 participants were excluded due to prior experience with mental illness ($n = 22$) or not meeting the minimum age requirement ($n = 2$). Therefore, 44 datasets were analyzed. The mean age of the participants was 20.86 years ($SD = 2.11$). The majority of 43 individuals self-indicated being female (age: $M = 20.86$, $SD = 2.13$), whereas only one person self-identified as male (age: 21). No one self-identified as a diverse gender.

4.2.4 Results and discussion

The average level of transportation into the story was 4.91 ($SD = 1.34$). This value is comparable to those achieved in related studies investigating transportation in written texts^[135,138,139], suggesting that our generated story elicited a sufficient degree of transportation.

For the three dimensions of state empathy, mean values of 2.00 ($SD = 0.86$) for *Affective Empathy*, 2.18 ($SD = 0.85$) for *Cognitive Empathy*, and 1.92 ($SD = 0.87$) for *Associative Empathy* were computed. As the scale has yet mainly used in combination with written news articles, video stimuli^[140,141], or immersive technologies^[142], but not written narratives, no data for direct comparisons are available. However, as the values range around the scale's center and are descriptively comparable to related works^[140,142], they were interpreted as sufficient.

Only one person provided qualitative feedback in the comment section, noting that the story was "very well-written" and that "it was easy to imagine oneself in the scene and empathize with the events", reinforcing the suitability of our story. Based on the overall feedback, we utilized the story without further adjustments.

4.3 Implementation

We implemented the story using the *Pepper* robot, generating four versions with either one type, both types, or no type of non-speech sounds added to the storytelling, resulting in four versions in total.

4.3.1 Story annotation

The story was annotated by a single human rater, the story author, using a flexible tokenization. This approach was chosen to allow for emphasizing the underlying intention with the annotation. As possible labels, we utilized the eight inner emotions of Plutchik's *Wheel of Emotions*^[143], namely *ecstasy*, *admiration*, *terror*, *amazement*, *grief*, *loathing*, *rage*, and *vigilance*, plus a *neutral* label. In doing so, *terror* was assigned six times, whereas *admiration*, *ecstasy*, and *grief* were assigned three times each. The rest of the text was annotated as *neutral*.

4.3.2 Robot behavior

The story was implemented using *Choregraphe* version 2.5.10.7 (<http://doc.aldebaran.com/2-5/software/choregraphe/index.html>). The *Pepper* robot was chosen due to its popularity and widespread use^[144], its relatively low intensity of consequential sounds^[145], and its extensive use in previous robotic storytelling research^[50,51,53,125,146,147].

Concerning body language, emotional expressions matching Plutchik's emotions retrieved from related work^[148] were matched to the annotated labels (Figure 1a). To improve the naturalness of the storytelling, further beat gestures and iconic gestures were implemented. Examples are shown in Figure 1b.

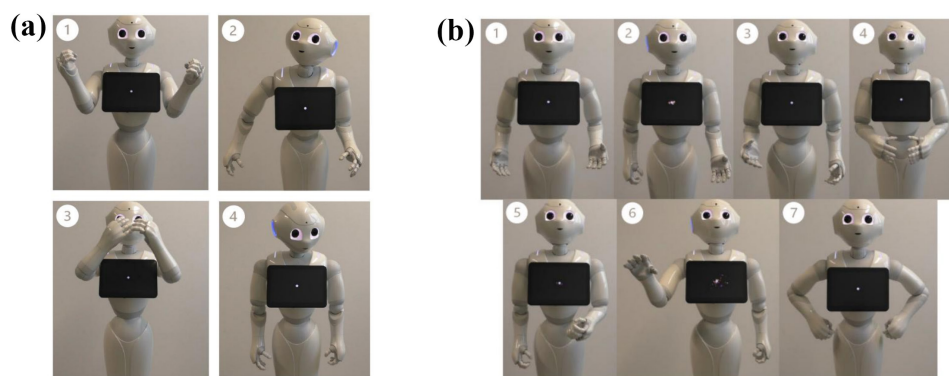


Figure 1. Robot's body language. (a) Emotional expressions for (1) ecstasy, (2) admiration, (3) terror; (b) Narrative behavior using the beat-gestures Explain-1 (1), and (4) grief. Explain-2 (2), Explain-4 (3), Explain-11 (4), Explain-5 (5), and iconic gestures No-1 (6) and YouKnowWhat-1 (7).

The auditory implementation of the story made use of the robot's preinstalled synthetic voice. The story alternates between the four perspectives of the storyteller, the protagonist's inner monologue, her spoken words, and the spoken words of a character met on the bus. The different perspectives were realized via differing pitch and speed of the robot's voice. The standard storyteller voice used a pitch of 85 percent and a speed of the standard 100 percent. The protagonist's inner monologues and spoken words used a pitch of 120 percent, while speed alternated between 100 and 115 percent based on her emotional state. For example, a fast and loud voice mirrored fear as it expresses a high state of emotional arousal^[149]. In doing so, the inner monologue was adjusted to represent emotional duress during the panic attack. Lastly, the voice of the boy on the bus was implemented with a pitch of 60 percent and the standard speed of 100 percent.

4.3.3 Music and sound effects

In order to underline the protagonist's emotional states, emotionally charged music was added to the story while sound effects were added to illustrate environmental cues.

The music was attached paragraph-wise based on the initial emotional annotation of the story. In accordance with Steinhäusser *et al.*^[53], four musical pieces were used, representing the emotions of *fear*—matched to *terror*, *joy*—matched to *ecstasy*, *sadness*—matched to *grief*, as well as *heroic* emotions, which were matched to *admiration*. The utilized pieces were “Torture Chamber” (<https://www.fesliyanstudios.com/royalty-free-music/download/torture-chamber/2856>) by David Robson, “Beautiful Memories” (<https://www.fesliyanstudios.com/royalty-free-music/download/beautiful-memories/234>), “A Sad Meme” (<https://www.fesliyanstudios.com/royalty-free-music/download/a-sad-meme/2479>) and “Inspirational Advertising” (<https://www.fesliyanstudios.com/royalty-free-music/download/inspirational-advertising-1/260>) by David Fesliyan, respectively.

The sound effects were individually assigned to the story text by the human annotator. They were retrieved from the platforms *Fesliyan Studios* (<https://www.fesliyanstudios.com/de/>) and *Freesound* (<https://freesound.org/>). Examples of sound effects are birds chirping, sirens of the rescue vehicle, or sounds representing the panic attack, such as a fast heartbeat. Sound effects representing the protagonist's breathing and laughing were recorded especially for this study by a young female voice actor. Some of the sound effects were combined using the tool *Audacity* (<https://www.audacityteam.org/>). For instance, to symbolize the protagonist's panic attack, the sound of a beating heart was combined with a high-pitched noise representing her ears ringing.

The sound effects used the standard pitch of 100 percent while the music was played at 50 percent in the background so as to not distract from the plot while still setting the scene.

In doing so, four conditions of the storytelling were created, either omitting (*Basic* condition) or integrating (*BothAdds* condition) both types of non-speech sounds, or integrating only background music (*Music* condition) or sound effects (*Effects* condition).

5. Methods

To evaluate the integration of non-speech sounds, namely background music and sound effects, into a robotic storytelling intervention to lower prejudice against people suffering from mental illnesses, a multi-methods laboratory study was conducted.

5.1 Study design

A 2 (integration or omission of music) $\times$ 2 (integration or omission of sound effects) between-subjects was applied, utilizing four versions of the robotic storytelling intervention as conditions: (1) storytelling integrating both background music and sound effects (*BothAdds*), storytelling integrating either (2) background music (*Music*) or (3) sound effects (*Effects*), and (4) a storytelling without any additional non-speech sounds (*Basic*).

5.2 Measures

We applied a multi-methods approach using both subjective self-reports of the participants as well as objective observations.

5.2.1 Observation of blinking behavior

Blinking behavior is strongly connected to the task that is currently performed^[150]. Since closing one's eyes, apparently, hinders information intake^[151], blinking is suppressed when acquiring visual information^[152] within, for example, difficult^[150,153] and attention-demanding tasks^[150]. Therefore, the blink rate adapts inversely to the attention spent on a task. This holds also true for non-visual tasks^[151] such as music listening^[154]. Based on this relation, we analyzed participants' blinking behavior as an objective measure to assess their *Attention* during the storytelling.

To measure blinking behavior, we video-recorded participants' faces during the storytelling intervention. The videos were then analyzed using a Python-implemented eye blink counter. As a base, we used the *face mesh module* and *plot module* from CVZone (<https://www.computervision.zone/courses/eye-blink-counter/>). We adapted the program to count blinks more accurately following the procedure described by Steinhäusser *et al.*^[53].

5.2.2 Self-reports via questionnaires

Recipients' storytelling experience was operationalized with transportation, using the same questionnaire as in the prestudy (TS-SF; see [Section 4.2.1](#)). Reliability computed for the current sample was .81.

Recipients' empathy with the story's protagonist was again assessed using the *State Empathy Scale* by Shen^[136] described in [Section 4.2.1](#). Cronbach's alpha for the current sample was .81 for *Affective Empathy*, .76 for *Cognitive Empathy*, .68 for *Associative Empathy*, and .88 for the whole scale.

Last, recipients' prejudice toward people with mental illness was measured using the Prejudice Towards People with Mental Illness scale by Kenny *et al.*^[66]. The scale comprises three subscales: (1) the *Fear/Avoidance* subscale including items on social distance and interaction difficulties (8 items, e.g., "It is best to avoid people who have mental illness."), (2) the *Malevolence* subscale with items on inferiority and unsympathetic attitudes (8 items, e.g., "People who develop mental illness are genetically inferior to other people."), (3) the *Authoritarianism* subscale on needs for control over affected people (6 items, e.g., "People who are mentally ill need to be controlled by any means necessary."), and (4) the *Unpredictability* subscale about unforeseeable behavior of affected persons (6 items, e.g., "People with mental illness often do unexpected things."). To adjust the answer format to the other questionnaires in the study, the original nine-point scale was adapted to a five-point scale, anchored by 1 - "Strongly disagree" and 5 - "Strongly agree". Kenny *et al.* report reliabilities of .91 for *Fear/Avoidance*, .80 for *Malevolence*, .79 for *Authoritarianism*, and .82 for *Unpredictability*. Cronbach's alpha for the current sample was .71 for *Fear/Avoidance*, .34 for *Malevolence*, .61 for *Authoritarianism*, and .72 for *Unpredictability*. Due to their low reliabilities, both the *Malevolence* and *Authoritarianism* subscales should be treated with caution.

5.3 Procedure and exclusion criteria

Once more, participants were acquired via the university's participant pool. In doing so, participants from the prestudy were excluded from participation in the main study. Similar to the prestudy, we implemented trigger warnings in the study advertisement. When arriving at the laboratory, participants first provided written informed consent to participate in the study.

Again, we integrated an aptitude test into our study protocol to exclude individuals with personal experience with OCD from the actual experiment. For this screening, again the personal relevance scale described in [Section 4.2.1](#) was utilized together with a question on prior OCD diagnosis, the minimum age requirement, and willingness to participate after reading the trigger warning

again. As the experimenter was aware of the test result in terms of inclusion or exclusion, we added an additional answer to the last question (not interested in the study), as well as an additional question on hearing impairment to make participants feel more comfortable if they get an exclusion notice in the presence of the experimenter. However, the experimenter had no access to the data collected in the aptitude test and was only informed about passing or failing the test.

After the aptitude test, the actual experiment started. Firstly, the participants received the story version to which they were randomly assigned. During the reception, their face was videotaped to allow for analyses of blinking behavior afterwards. Then, they filled in the questionnaires on transportation, state empathy, and prejudice against people suffering from mental illnesses. Next, participants provided demographic data (age, gender, prior experience with the robot) and were invited to leave a comment. Lastly, the participants were thanked, debriefed about the study goal, and handed a list of contact points such as the crisis hotline or the university's psychotherapeutic counseling.

5.4 Participants

In total, 131 individuals started the experiment, however, ten of them were excluded due to failing the aptitude screening. In consequence, 121 individuals with a mean age of 21.66 (SD = 3.14) took part in the study. The majority of 94 participants self-identified as female (age: M = 21.38, SD = 3.19), whereas 27 of them self-identified as male (age: M = 22.63, SD = 2.80). No one self-identified as a diverse gender.

Being randomly assigned to the conditions, 30 participants each received the *Music* (23 female, 7 male, 0 diverse; age: M = 21.27, SD = 2.35), *Effects* (22 female, 8 male, 0 diverse; age: M = 21.17, SD = 1.88), respectively *Basic* condition (24 female, 6 male, 0 diverse; age: M = 22.50, SD = 4.65), whereas 31 individuals received the *BothAdds* condition (25 female, 6 male, 0 diverse; age: M = 21.71, SD = 2.95).

6. Results

All analyses were carried out using JASP version 0.19.1.0^[155] and a significance level of .05. Descriptive values are displayed in [Table 1](#).

Table 1. Descriptive data from main study.

	<i>BothAdds</i>		<i>Music</i>		<i>Effects</i>		<i>Basic</i>	
	M	SD	M	SD	M	SD	M	SD
Transportation ^a	5.52	0.82	4.71	1.24	4.71	1.27	5.21	1.03
Affective Empathy ^b	2.14	0.88	1.68	0.81	1.90	0.74	1.90	0.79
Cognitive Empathy ^b	2.21	0.81	1.90	0.86	2.23	0.86	2.02	0.63
Associative Empathy ^b	2.02 ^A	0.63	1.42 ^{A,B}	0.76	1.68	0.68	1.93 ^B	0.70
Fear/Avoidance ^c	1.87	0.50	1.86	0.49	1.91	0.59	1.91	0.52
Malevolence ^c	1.44	0.32	1.36	0.23	1.40	0.33	1.38	0.30
Authoritarianism ^c	1.65	0.38	1.86	0.54	1.82	0.58	1.61	0.44
Unpredictability ^c	3.00	0.62	2.83	0.58	2.95	0.48	3.12	0.70
Blink Rates ^d	263.00	109.28	226.33	113.45	287.93	146.90	276.47	131.67

Basic: condition without additional non-speech sounds; *Effects*: condition with sound effects; *Music*: condition with background music; *BothAdds*: condition with sound effects and background music. Means within lines with identical superscripts (A, B) differ significantly at least $p < .05$ as calculated by Bonferroni-corrected *post hoc* tests. ^a: calculated values from 1 to 7; ^b: calculated values from 0 to 4; ^c: calculated values from 1 to 5; ^d: counted during the whole storytelling as explained in [Section 5.2.1](#); M: mean; SD: standard deviation.

6.1 Blink rates

Each of the 121 trials was video-recorded. Unfortunately, one video of the *BothAdds* condition was corrupted and thus excluded from the analysis, resulting in 30 videos per condition. Out of the 120 valid video recordings, outliers as well as randomly chosen ones were cross-coded by a human rater for quality assurance, resulting in cross-coding of 13 videos. For these videos, the human rater's counts were adopted. The interrater reliability between the program and both human raters was Krippendorff's $\alpha = .63$, calculated using the *k-alpha calculator*^[156]. Blink rates per condition are illustrated in [Figure 2](#).

Shapiro Wilk tests indicated a normal distribution ($ps > .5$). The calculated Levene's test validated equality of variances ($p = .592$). Again, a 2 (integration or omission of music) $\times$ 2 (integration or omission of sound effects) analysis of variance (ANOVA) was computed

to investigate differences in blinking behavior between the conditions. Results indicated no significant main effect of music or sound effects, nor a significant interaction effect. A planned contrast comparing the *Basic* condition to the conditions integrating non-speech sounds yielded no significant difference ($t(116) = -0.65, p = .515, d = .41$). Similarly, the planned contrast comparing the *BothAdds* condition to the ones integrating only one type of non-speech sound indicated no significant difference ($t(116) = -0.21, p = .836, d = .09$).

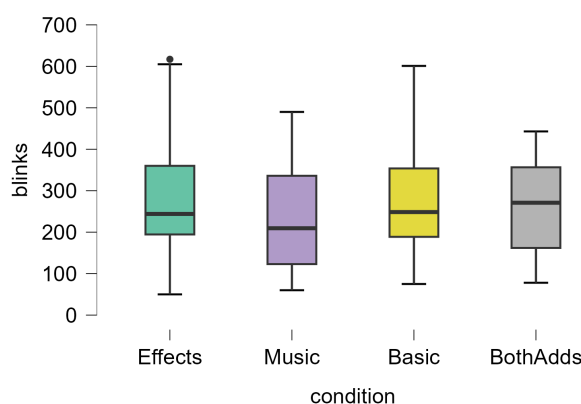


Figure 2. Blink count per condition. Error bars represent standard errors.

6.2 Questionnaire data

For all variables, 2 (integration or omission of music) $\times$ 2 (integration or omission of sound effects) comparisons were computed. Their results are presented in [Table 2](#). If applicable, Bonferroni-Holm corrected *post hoc* pairwise comparisons were carried out. To gain deeper insights into the added value of non-speech sounds in general compared to their omission as well as the combination of non-speech sounds compared to using single types only, contrasts were specified (*Basic* vs. rest; *BothAdds* vs. *Music* and *Effects*).

Table 2. ANOVA results for questionnaire data and blink rates.

	Main Effect Music			Main Effect Sound Effects			Interaction Effect		
	<i>F</i>	<i>p</i>	ω^2	<i>F</i>	<i>p</i>	ω^2	<i>F</i>	<i>p</i>	ω^2
Transportation	0.60	.442	.00	0.60	.442	.00	10.65	.074	.01
Affective Empathy	0.01	.945	.00	2.39	.125	.01	2.39	.125	.01
Cognitive Empathy	2.01	.159	.01	0.53	.468	.00	1.60	.208	.01
Associative Empathy	0.44	.508	.00	1.91	.169	.01	11.77	< .001***	.08
Fear/Avoidance	0.24	.623	.00	0.00	.988	.00	0.00	.948	.00
Malevolence ¹	0.01	.913	.00	0.77	.382	.00	0.32	.571	.00
Authoritarianism ²	0.26	.612	.00	0.00	.999	.00	5.64	.019*	.04
Unpredictability	1.15	.286	.00	0.00	.956	.00	2.47	.119	.00
Blink Rates	2.65	.106	.01	1.09	.299	.00	0.30	.586	.00

¹: presented for comparison purposes only; interpretability is limited due to unacceptable reliability; ²: should be interpreted with caution due to questionable reliability. * $p < .05$, ** $p < .01$, *** $p < .001$. ANOVA: analysis of variance.

6.2.1 Storytelling experience

For *Transportation*, homogeneity of variances was suggested by Levene's test ($p = .179$). Shapiro-Wilk tests indicated a violation of the normality assumption for the *Music* condition ($p < .001$). However, given the reported robustness of ANOVA against violations of normality^[157], two-way ANOVAs were performed to examine the main effects and interactions of non-speech sound additions to the robotic storytelling. No significant main effects of music nor sound effects were obtained. Similarly, no significant interaction was found. Planned contrasts indicated no significant difference between the conditions with non-speech sounds added compared to the *Basic* condition ($t(117) = -0.99, p = .324, d = .63$) but a significant difference between the conditions with only one type of non-speech sound compared to the *BothAdds* condition ($t(117) = -3.32, p = .001, d = 1.47$) with higher values for the combined non-speech sounds condition.

6.2.2 Empathy

As state empathy was assessed on three dimensions, a multivariate ANOVA (MANOVA) was computed. A Shapiro-Wilk test indicated multivariate normality ($p = .168$). Levene's tests indicated homogeneity of variances for all dimensions ($ps > .05$). Correlations between dependent variables were low ($rs < .73$), indicating that multicollinearity was not a confounding factor in the analysis. The calculated MANOVA did not show statistically significant main effects of music (Pillai's Trace = .03, $F(3,115) = 0.97$, $p = .409$) or sound effects (Pillai's Trace = .02, $F(3,115) = 0.95$, $p = .421$) on the combined dependent variables, but a significant interaction of both (Pillai's Trace = .12, $F(3,115) = 5.19$, $p = .002$). *Post-hoc* univariate ANOVAs were conducted for every empathy dimension.

Regarding *Affective Empathy*, the follow-up univariate ANOVA yielded no significant main effect of music or sound effects. Similarly, the interaction effect was not significant. The planned contrasts comparing the *Basic* to the three other conditions ($t(117) = 0.04$, $p = .968$, $d = .03$) as well as the one contrasting the *BothAdds* to the *Music* and *Effects* conditions ($t(117) = -1.94$, $p = .055$, $d = .86$) also revealed no significant differences.

Likewise, no significant main effect of music or sound effects was found for *Cognitive Empathy*. Moreover, the interaction effect was not significant. Neither the planned contrast comparing the *Basic* condition to the ones with additional non-speech sounds ($t(117) = -1.13$, $p = .263$, $d = .71$) nor the contrast between the conditions with only one additional sound type and the *BothAdds* condition ($t(117) = -0.76$, $p = .450$, $d = .34$) revealed significant differences.

Lastly, for *Associative Empathy*, again no significant main effects of music or sound effects were found. However, a significant interaction effect was obtained. Pairwise *post-hoc* comparisons indicated significantly lower values in the *Music* compared to the *Basic* ($p = .028$, $d = .75$) as well as compared to the *BothAdds* condition ($p = .005$, $d = .88$). While the planned contrast between the *Basic* and the three other conditions yielded no significant difference ($t(117) = -1.56$, $p = .121$, $d = .99$), the difference in *Associative Empathy* between the *BothAdds* condition and the conditions integrating only one additional sound type indicated significantly higher values in the *BothAdds* condition ($t(117) = -3.12$, $p = .002$, $d = 1.38$).

6.2.3 Prejudice

Since prejudice was measured on four dimensions, a MANOVA was computed. A Shapiro-Wilk test indicated lack of multivariate normality ($p < .001$). However, Finch^[158] reported robustness of the MANOVA against violation of the normality assumption. Levene's tests indicated homogeneity of variances for all dimensions ($ps > .05$). Correlations between dependent variables were low ($rs < .33$), indicating that multicollinearity was not a confounding factor in the analysis. The calculated MANOVA did not show statistically significant main effects of music (Pillai's Trace = .02, $F(4,114) = 0.44$, $p = .777$) or sound effects (Pillai's Trace = .01, $F(4,114) = 0.21$, $p = .933$) on the combined dependent variables. In contrast, a significant interaction of both (Pillai's Trace = .09, $F(4,114) = 2.69$, $p = .035$) was indicated. Therefore, *post-hoc* univariate ANOVAs were conducted for every prejudice dimension. However, due to the low reliability of the subscales *Malevolence* and *Authoritarianism* reported in Section 5.2.2, the respective results should be treated with caution and are presented separately.

Regarding self-reported *Fear/Avoidance* behavior, no significant main effect of music or sound effects, nor a significant interaction effect was obtained. Similarly, planned contrasts did not indicate significant differences between conditions with non-speech sounds and the *Basic* condition ($t(117) = -0.33$, $p = .742$, $d = .21$), nor between the *BothAdds* condition and those integrating only one type of non-speech sounds ($t(117) = 0.16$, $p = .877$, $d = .07$).

Concerning *Unpredictability*, again no significant main effects, neither of music nor of sound effects, or an interaction effect were obtained. Similarly, planned contrasts indicated no significant difference between the *Basic* and the other conditions ($t(117) = -1.55$, $p = .123$, $d = .98$), as well as contrasting the *BothAdds* condition to the conditions integrating either music or sound effects ($t(117) = -0.83$, $p = .409$, $d = .37$).

Although results should be treated with caution due to low reliability, we computed the same tests for the subscales *Malevolence* and *Authoritarianism*. This again resulted in insignificant group differences, but a significant interaction effect for *Authoritarianism*. However, pairwise *post hoc* comparisons yielded no significant differences between the individual conditions. Regarding planned contrasts, for *Malevolence*, again no significant difference between the *Basic* and the other conditions ($t(117) = 0.24$, $p = .810$, $d = .15$), as well as contrasting the *BothAdds* condition to the conditions integrating either music or sound effects ($t(117) = -0.87$, $p = .385$, $d = .39$) were indicated. For *Authoritarianism*, the planned contrasts also revealed no significant differences between the *Basic* and the three other conditions ($t(117) = 1.66$, $p = .099$, $d = 1.05$) and the *BothAdds* compared to both the *Music* and *Effects* ($t(117) = 1.74$, $p = .084$, $d = .77$) conditions.

6.2.4 Transportation as a mediator

To analyze the mediating effect of participants' transportation into the story on their empathy toward the story protagonist, we computed a mediation analysis with bootstrapping with 5,000 bootstrap samples. Music and sound effects (present vs. absent) served as predictors, transportation as mediator, and affective, cognitive, and *associative empathy* as outcome variables. The estimated

model is depicted in Figure 3. As indicated by R^2 values, the model explained only 1.0% of the variance in *Transportation*, but 46.5% of the variance of *Affective*, 22.3% of the variance of *Cognitive*, and 43.0% of the variance of *Associative Empathy*. We did not observe any significant total effects ($ps > .05$). Moreover, no significant indirect effects of music or sound effects on any empathy dimension via *Transportation* were revealed ($ps > .05$), indicating no mediating role of transportation. In contrast, a significant direct effect of music on *Cognitive Empathy* ($p = .044$) was indicated, whereas no direct effects emerged for music on *Affective Empathy* ($p = .533$) or *Associative Empathy* ($p = .138$). For sound effects, no direct effect was found for *Cognitive* ($p = .658$), *Affective* ($p = .188$) or *Associative Empathy* ($p = .269$). Regarding the path coefficients, neither music ($p = .441$) nor sound effects ($p = .441$) predicted *Transportation* significantly, while *Transportation* significantly predicted *Affective* ($p < .001$), *Cognitive* ($p < .001$), and *Associative Empathy* ($p < .001$).

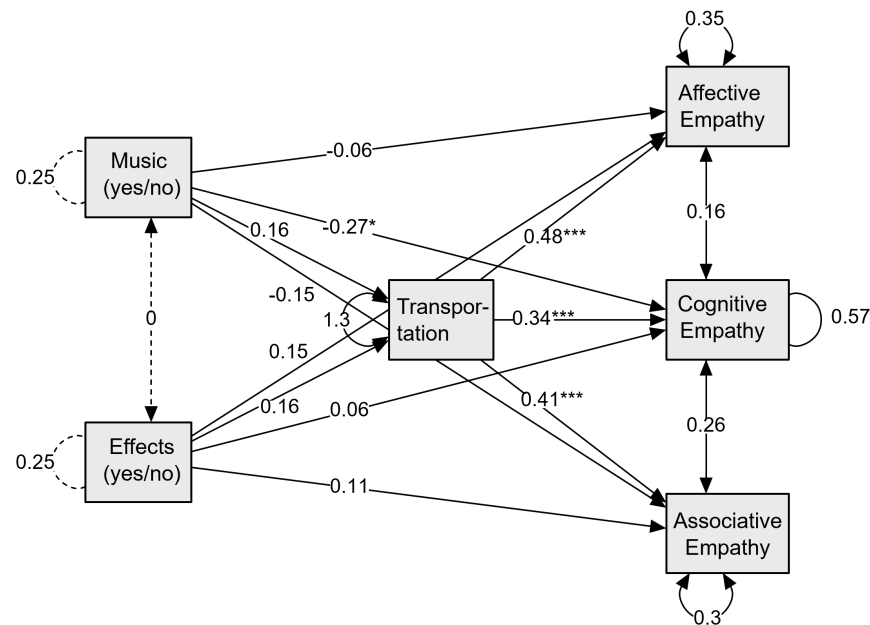


Figure 3. Mediation model predicting empathy.

To investigate the mediating effect of participants' transportation on their prejudice toward individuals affected by mental disorders measured after the intervention, we computed a mediation analysis using the presence or absence of music and sound effects as predictors, transportation as mediator, and the prejudice dimensions *Fear/Avoidance*, *Malevolence*, *Authoritarianism* and *Unpredictability* as outcome. Again, 5,000 bootstrap samples were used. The estimated model is depicted in Figure 4. As indicated by R^2 values, the model explained only small proportions of variance in prejudice, namely 1% of the variance of *Unpredictability*, 3.5% of the variance for *Authoritarianism*, 4.6% of the variance for *Fear/Avoidance*, and 7.3% of the variance for *Malevolence*. We did not observe any significant total effects ($ps > .05$). Furthermore, no significant indirect effects of music or sound effects on any prejudice subscale through *Transportation* were revealed ($ps > .05$), indicating that there was no mediation. Also, no significant direct effects were observed ($ps > .05$). Regarding the path coefficients, neither music ($p = .447$) nor sound effects ($p = .444$) significantly predicted *Transportation*. In contrast, *Transportation* significantly predicted the prejudice dimensions *Fear/Avoidance* ($p = .034$), *Authoritarianism* ($p = .042$), and *Malevolence* ($p = .005$) but not *Unpredictability* ($p = .769$). Again, reliabilities for the *Malevolence* and *Authoritarianism* subscales were low, thus interpretability is limited.

6.3 Qualitative data

Only 19 participants provided qualitative feedback using the comments section. Seven comments were entirely positive ($n_{BothAdds} = 3$, $n_{Music} = 2$, $n_{Effects} = 1$, $n_{Basic} = 1$), while ten comments were of negative valence ($n_{BothAdds} = 1$, $n_{Music} = 5$, $n_{Effects} = 3$, $n_{Basic} = 1$). Lastly, two comments were neutral ($n_{Effects} = 1$) or mixed in their valence ($n_{BothAdds} = 1$). Concerning positive feedback, participants liked the overall study idea of a robotic storytelling on mental health issues ($n_{BothAdds} = 2$, $n_{Music} = 2$, $n_{Effects} = 1$), e.g., "The story was very interesting to hear and really moved me. I would be delighted to participate in further studies with Pepper and mental illness in the future!", and enjoyed listening to the robot ($n_{Music} = 1$, $n_{Basic} = 1$). One person in the *BothAdds* condition specifically referred to the music and sound effects as "effective", whereas another person from this group expressed specific liking for the robot's body language. In contrast, other participants claimed the body language to be distracting ($n_{Music} = 2$) and the robot's voice to be hard to understand ($n_{Music} = 1$, $n_{Effects} = 1$, $n_{Basic} = 1$), e.g., "The robot spoke quite quickly and somewhat unclearly, and I didn't understand some of the important words", also due to the sound effects ($n_{Effects} = 1$). One person found the music partly inappropriate ($n_{Music} = 1$). Lastly, three participants found it difficult to answer the questionnaires ($n_{Music} = 1$, $n_{Effects} = 2$).

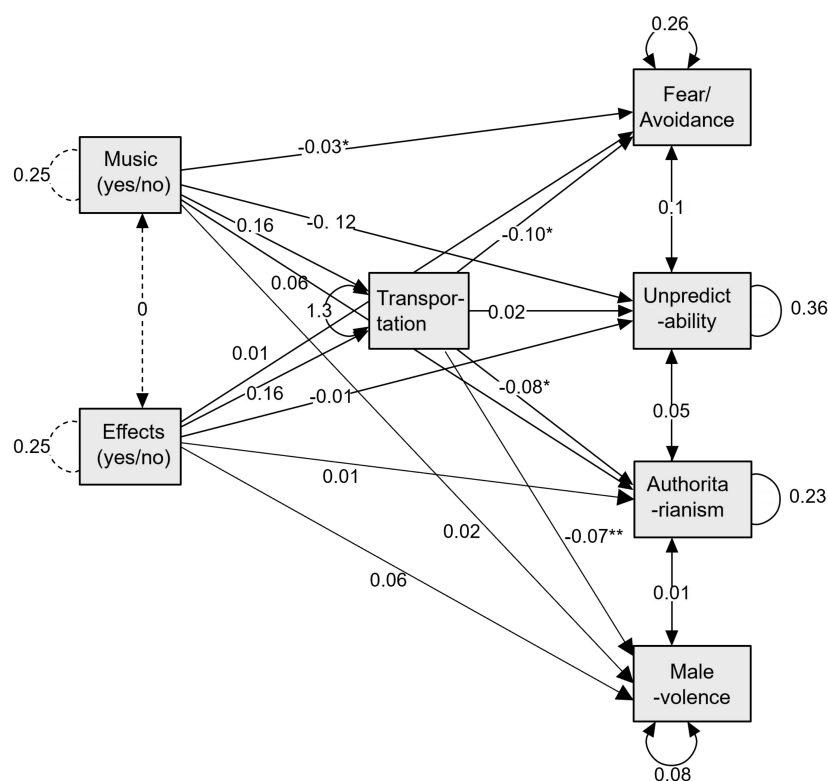


Figure 4. Mediation model predicting prejudice.

6.4 Comparison to written story

As an exploratory analysis, we compared the storytelling experience between participants reading the story in the prestudy and participants receiving the story from the robot in the main study. Prior to statistical testing, assumptions were checked. Levene's tests indicated equality of variances for all variables, but Shapiro-Wilk tests revealed a violation of the normality assumption for all variables ($ps < .05$). Thus, non-parametric Mann-Whitney U-tests were conducted comparing *Transportation* as well as *Empathy* between the prestudy and the main study regardless of the sound conditions.

As shown in Table 3, no significant differences between the text version and the robotic storytelling were obtained. Descriptively, *Transportation* was slightly higher in the main ($M = 5.04$, $SD = 1.14$) than in the prestudy ($M = 4.91$, $SD = 1.34$), whereas *Affective*, *Cognitive* and *Associative Empathy* scored slightly higher in the prestudy ($M_{Affective} = 2.00$, $SD_{Affective} = 0.86$; $M_{Cognitive} = 2.18$, $SD_{Cognitive} = 0.85$; $M_{Associative} = 1.97$, $SD_{Associative} = 0.87$) compared to the main study ($M_{Affective} = 1.91$, $SD_{Affective} = 0.81$; $M_{Cognitive} = 2.17$, $SD_{Cognitive} = 0.86$; $M_{Associative} = 1.76$, $SD_{Associative} = 0.87$).

Table 3. Test results for the comparison between prestudy and main study ($N = 165$).

	<i>U</i>	<i>p</i>	<i>r_{rb}</i>
Transportation	2,743.00	.766	.03
Affective Empathy	2,472.50	.484	.07
Cognitive Empathy	2,701.00	.887	.00
Associative Empathy	2,287.50	.166	.14

7. Discussion

We conducted a laboratory study to investigate the beneficial effects of a robotic storyteller's multimodality, focusing on its ability to seamlessly integrate music and sound effects, on its deployment as an anti-stigma intervention. In doing so, we compared storytelling experience, emotion induction, attention, and prejudice measured after the intervention. The high number of participants excluded from the experiment, i.e., due to their own experiences with mental illness, 24 in the pretest and 10 in the main study, underpin the importance of the topic.

As indicated by participants' blink rates, no differences in attention during story reception were observed between the sound conditions. Thus, **H1a** and **H1b** are both rejected. Although listening to music has been shown to decrease blink rates compared to silence^[154], suggesting heightened attention, our results indicate that neither background music nor sound effects or their combination provided an additional attentional benefit beyond the attention elicited by the storytelling itself. This finding is quite surprising given the positive correlation reported by Santavirta *et al.* between audio intensity and blink rates in movie watching analyses^[159]. However, film audio is closely connected to both visuals and the narrative, and was not manipulated in their experiment. Comparing blink rates of participants viewing different versions of a movie with the same narrative and content, Andreu-Sanchez *et al.*^[160] found that editing style did not influence the timing of blinking, suggesting that rather narrative content than visual style is decisive for when to blink. The same may hold true for auditory style. Future work across different storytelling media is needed to shed more light on the relationship between non-speech sounds and attentional blinking in mediated storytelling, including robotic storytellers. Ideally, such studies should incorporate additional measures, such as questionnaires, to provide a more comprehensive understanding of these effects.

We found no significant differences between the conditions integrating additional sounds and the control condition concerning recipients' transportation into the story, resulting in the rejection of **H2a**. This finding is in line with related research examining the integration of music and sound effects in robotic storytelling across different story genres^[53,128]. In contrast to related work recommending the use of music only^[53], we observed a positive effect of combined versus single sound additions, resulting in stronger narrative transportation for recipients in the *BothAdds* condition. Accordingly, **H2b** is accepted. The reason could be the utilized sound effects. While^[53] employed romantic short stories that focus on interactions between (to-be) couples and added sound effects primarily to illustrate the environment, our story centered on the experience of a single individual, using sound effects not only to depict the environment but also to convey the protagonist's inner thoughts and emotions, for example through rapid breathing. Although we did not assess participants' perceived appropriateness of the sounds, employing the full interpretive range of sound effects to communicate both descriptive^[60,94,100] and emotional information^[98,101], the stimuli integrating sound effects in our study, namely the conditions *Effects* and *BothAdds*, may have been better suited to provide insight into the transportative potential of sound effects, as reflected in the higher values compared to^[53]. Therefore, in terms of improving storytelling experience, we not only recommend the combination of sound effects and background music for robotic storytelling, but also the integration of sound effects to convey both environmental and emotional information. Furthermore, compared to related works employing relatively short sound effects, such as "the sound of something that is dropped on the floor, swallow and rustling paper, or opening a door"^[53], we used longer-lasting sound effects, such as breathing or heartbeats, facilitating the appropriate time-wise placement of the sound effects. This way, the timing might have been improved, highlighting the need for future research to systematically explore not only the type of sound effects integrated, but also their timing to achieve generalizability. Interestingly, comparing our transportation values to the same related work^[53] as well as further related studies^[128,161], we find generally improved transportation across all conditions in our study. This finding suggests the appropriateness of the story theme for robotic storytelling.

In contrast to our findings on transportation, combining both types of non-speech sounds did not enhance participants' empathy towards the story protagonist, leading to the rejection of **H3**. However, while the integrated non-speech sounds did not influence participants' *Affective Empathy*, their affective reactions to the protagonist's emotions^[136], nor their *Cognitive Empathy*, their ability to adopt the protagonist's perspective^[136], they did influence their *Associative Empathy*, reflecting participants' sense of "social bonding" with the protagonist^[162]. We found a significant advantage of both the *Basic* and the *BothAdds* conditions over the *Music* condition for *Associative Empathy* as well as a significant difference contrasting the *BothAdds* condition to those integrating only one type of non-speech sounds (**RQ1**). Music seemed to increase the psychological distance between recipients and the story protagonist. Our mediation analysis sheds light on the underlying mechanisms. Using the presence or absence of music and sound effects as a predictors, we found that *Transportation* does not mediate the effect of the experimental manipulation on empathy. Nevertheless, *Transportation* appears to be a significant predictor of all three dimensions of empathy, while being only minimally influenced by the integration of non-speech sounds itself. The effect of music integration on *Associative Empathy* did not reach significance in the mediation model, likely due to opposing patterns in the *Music* and *BothAdds* conditions, which seem to cancel each other out. In contrast, a significant direct effect of music on *Cognitive Empathy* emerged, which only manifested when controlling for *Transportation*. This suggests that, although *Transportation* positively predicts *Cognitive Empathy*, the integration of music itself may have a negative influence, which is suppressed by the effect of *Transportation*. As a result, when controlling for *Transportation*, the integration of music was associated with slightly lower levels of *Cognitive Empathy*. As with *Associative Empathy* in the ANOVA analysis, music may have shifted attention away from cognitive perspective-taking processes. This result contradicts findings reported by McDonald *et al.*^[163], who observed that sad music accompanying videos of emotional distress in others improved empathy compared to neutral music. In our study, however, the story concluded with a happy ending, and emotionally-congruent heroic music was used to illustrate the protagonist's pride in managing her chronic condition. The rapid change in emotional valence may have caused confusion, potentially negatively influencing the positive effect of music on empathy. Taken together, these mediation analysis results suggest that *Transportation* seems to be a competing predictor rather than a mediator of music integration in our study. Music only improved empathy when presented together with sound effects, strengthening the recommendation to combine both types of non-speech sounds in robotic storytelling.

Concerning prejudice toward individuals with mental illness, the manipulation of non-speech sounds did not yield any significant effects. While the results concerning *Malevolence* are not interpretable due to unacceptable internal consistency, none of the four dimensions, namely *Fear/Avoidance* behavior, *Malevolence*, *Authoritarianism*, or perceived *Unpredictability*, differed between the conditions, neither when contrasting non-speech sound integration with *Basic* nor when comparing single sound types to *BothAdds*. Accordingly, *H4a* and *H4b* were rejected. These null findings may partly be explained by floor effects, which also clarify the low reliabilities computed. For all dimensions except *Unpredictability*, mean scores were low and showed limited variance, leaving little room for between-condition differences. Mental health issues are particularly prevalent among university students compared to the general population of the same age^[164]. Although individuals with personal or close OCD experience were excluded, participants may still have had greater contact with affected individuals than the broader population, potentially contributing to generally low prejudice levels. Furthermore, particularly the *Malevolence* and *Authoritarian* subscales are referred to as “unfavourable attitudes”^[13], explaining the drop in internal consistency. Future evaluations should integrate measures on social desirability to account for such answer patterns. Moreover, given the generally low levels observed, incorporating pre-intervention assessments of prejudice or a placebo condition, such as a robotic storytelling of another genre, in future studies would allow for a more accurate evaluation of potential changes. Compared to the other dimensions, *Unpredictability* is more subtle^[13], presumably explaining the absence of floor effects in this dimension. In line, related studies report higher values on this compared to the other dimensions^[66,165]. One possible explanation for the absence of experimental effects is that *Unpredictability* represents a more cognitively entrenched dimension^[166,167]. Although our story conveyed self-efficacy of the protagonist and illustrated her control over the situation across all four conditions, our sound manipulation rather targeted emotional processing of the story. Thus, it may not have added persuasive impact beyond the information already conveyed in the story, resulting in comparable prejudice levels in terms of *Unpredictability* across conditions. A similar picture is drawn by our mediation results. Although *Transportation* did not mediate the sound conditions and prejudice dimensions, leading to the rejection of *H5b*, *Transportation* significantly predicted all prejudice dimensions except for *Unpredictability*. This finding strengthens the suggestion that this cognitive facet of stigma is less influenced by storytelling experience and its manipulations, than the more affective and evaluative aspects of prejudice are. However, the small proportion of explained variance in our model suggests that although *Transportation* affects prejudice, its influence is limited, indicating that other factors likely play a substantial role.

Comparing the results of the main study with those from the online prestudy revealed no significant differences in *Transportation* or empathy. This suggests that, despite participants in the main study reporting that the robot’s voice was difficult to understand, the overall processing fluency did not appear to hinder transportation^[168,169]. This finding is encouraging, as traditional stigma-reduction interventions frequently rely on written materials such as vignettes^[15,170], but do not bear the opportunity to add further modalities. Robotic storytelling may therefore represent an equally effective, or potentially even more effective, alternative. Nevertheless, future research should employ pre-post designs and/or placebo conditions to more precisely assess intervention-related changes.

8. Future Work

Although the sound manipulation did not yield the expected effects, our findings suggest that transportation may represent a key mechanism for enhancing the effectiveness of stigma-reducing robotic storytelling interventions. Strengthening transportation may therefore improve overall intervention outcomes. Robotic modalities that have been shown to foster transportation, such as facial expressions^[43,119], should be systematically examined in future research with regard to their potential to reduce prejudice. In line, the usage of different robot models, not only being capable of facial expressions but also differing in size or voice, should be taken into account in future studies to allow for greater generalizability of the findings. Moreover, Steinhäusser *et al.*^[45] reported positive effects of interactivity in a robotic storytelling intervention against bullying for children, particularly in increasing bullying awareness. This suggests that interactivity may also be a promising component in stigma-reducing robotic storytelling and should be considered in future intervention designs. Finally, findings from chatbot research suggest that a first-person narrative perspective is more effective in simulating social contact and reducing mental health stigma than a third-person perspective^[15]. Accordingly, future research should also consider manipulations of the narrative itself, such as perspective, as potential mechanisms for enhancing stigma-reduction effects.

While we primarily investigated the role of transportation affecting intervention outcomes, namely empathy and prejudice, in line with our focus on storytelling, Birtel and Oldfield^[167] examined empathy itself as a mediating variable, reporting a significant negative relationship with stigma-relevant attitudes. In addition, they identified intergroup anxiety as an important mediator, which was not assessed in our study. Pettigrew and Tropp^[16] suggest that perhaps anxiety reduction is a prerequisite to increasing empathy and perspective taking for prejudice reduction. Consequently, the variable of intergroup anxiety should be taken into account in future studies to examine the full potential of robotic storytelling as an intervention tool to decrease mental health prejudice.

Prejudice also varies across specific mental disorders. For example, stigma tends to be stronger toward schizophrenia than toward depression^[68], at least in Western countries^[171]. To our knowledge, however, comparable research examining OCD in this context is scarce, although existing evidence indicates substantial stigma toward individuals with OCD^[172], particularly when symptoms involve taboo or unacceptable thoughts, such as those related to sexuality^[173], as depicted in our narrative. The specific form of OCD

addressed in our study, namely pedophilia-themed OCD, may have introduced additional bias due to its association with a highly stigmatized and criminalized topic, potentially limiting our findings. The purpose of the story was to illustrate that OCD-related intrusive thoughts can be entirely inconsistent with an individual's actual personality, values, and intentions. However, future studies on robotic interventions against stigma toward mental illness should examine different forms of mental illness and OCD symptom presentations in order to compare the effectiveness between them. Furthermore, the study was limited by its sample, which consisted predominantly of young female participants, as well as by its design, which assessed stigma only after the intervention. Consequently, changes in prejudice could not be evaluated relative to baseline levels. Therefore, future studies should not only integrate a placebo condition and a pre-measurement of prejudice to gain better insights into the stigma change, but also test the robotic storytelling intervention with a more diverse participant sample to improve the generalizability of the findings. Moreover, future research should adopt longitudinal designs with repeated interventions, given that the effects of single exposure interventions have been shown to fade as stereotypes re-emerge over time^[71]. In doing so, the suitability of social robots for stigma-reducing storytelling interventions suggested by our results could be further explored in the future.

9. Conclusion

Despite greater public awareness, stigma against people with mental illness is still a serious issue in today's society. It impairs the lives of those affected by undermining their self-confidence and well-being. Storytelling-based anti-stigma interventions can help reduce prejudice by promoting education and substituting for direct contact. Social robots as physically embodied agents capable of multimodal communication offer additional advantages for such interventions. In particular, the integration of music and sound effects may benefit the outcomes. To this end, we evaluated a robotic storyteller as an intervention method, focusing on the influence of additional sound integration, namely background music and sound effects. While the integration of non-speech sounds did not directly reduce prejudice, their combined use enhanced transportation and associative empathy, highlighting the importance of carefully designed multimodal storytelling. In contrast, music presented in isolation yielded less favorable effects, suggesting that isolated auditory augmentation may not be sufficient and could interfere with certain empathic processes. Moreover, transportation emerged as a central mechanism, predicting empathy and, in part, prejudice-related outcomes. Although mediation effects were not supported, the results underscore transportation as a promising lever for strengthening the impact of robotic anti-stigma interventions. Enhancing recipients' transportation into the story may therefore be more critical than simply increasing sensory richness used to convey emotions.

In summary, while our study provided valuable insight into the potential of robotic storytellers in mental health stigma interventions, and particularly into the transportation they can foster, future research should investigate how this transportation can be systematically stimulated through narrative perspective, interactivity, or additional multimodal cues. Moreover, future studies should employ longitudinal pre-post designs and placebo conditions to better capture changes in prejudice over time. Such work will be essential for advancing robotic storytelling as an evidence-based tool to reduce mental health stigma.

Supplementary materials

The supplementary material for this article is available at: [Supplementary materials](#).

Declarations

Acknowledgements

The authors would like to thank Sandra Molter from the Contact and Information Center for Students with Disabilities and Chronic Illnesses (KIS) at the University of Würzburg for her valuable advice on participant recruitment and in developing the story used in this study. The authors also thank Sophia Maier and Ohenewa Bediako Akuffo for their help in annotating the observational data.

Authors contribution

Steinhaeuser SC: Conceptualization, methodology, data curation, formal analysis, resources, visualization, writing-original draft, writing-review & editing.

Semineth M: Conceptualization, methodology, investigation, writing-original draft.

Lugrin B: Conceptualization, supervision, writing-review & editing.

Conflicts of interest

The authors declare no conflicts of interest.

Ethical approval

The presented study was approved by the ethics committee of the Human-Computer-Media Institute at the University of Würzburg

(vote #240823).

Consent to participate

Written informed consent to participate in the study was obtained from all participants.

Consent for publication

Not applicable.

Availability of data and materials

Data supporting the findings of this study are available from supplementary materials and the corresponding authors upon reasonable request.

Funding

None.

Copyright

© The Author(s) 2026.

References

1. Mental disorders [Internet]. Geneva: World Health Organization; 2025. Available from: <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>
2. Corrigan PW, Green A, Lundin R, Kubiak MA, Penn DL. Familiarity with and social distance from people who have serious mental illness. *Psychiatr Serv*. 2001;52(7):953-958. [DOI][PubMed]
3. Mental health [Internet]. Hamburg: Statista; 2025. Available from: <https://de.statista.com/statistik/studie/id/119901/dokument/mental-health/>
4. Depression: Stigma and mental health [Internet]. Oxford: University of Oxford; 2004. Available from: <https://www.hexi.ox.ac.uk>
5. Mental Health: Market data & analysis [Internet]. Hamburg: Statista GmbH; 2025. Available from: <https://www.statista.com/study/119816>
6. Jorm AF, Korten AE, Jacomb PA, Christensen H, Rodgers B, Pollitt P. "Mental health literacy": A survey of the public's ability to recognise mental disorders and their beliefs about the effectiveness of treatment. *Med J Aust*. 1997;166(4):182-186. [DOI][PubMed]
7. Mental illness: Estimated cases, World, 2021 [Internet]. Seattle: IHME; 2024. Available from: <https://ourworldindata.org/grapher/mental-illness-estimated-cases?time=2021>
8. Corrigan PW, Markowitz FE, Watson AC. Structural levels of mental illness stigma and discrimination. *Schizophr Bull*. 2004;30(3):481-491. [DOI]
9. Corrigan PW, Bink AB. The stigma of mental illness. In: *Encyclopedia of mental health*. Amsterdam: Elsevier; 2016. p. 230-234. [DOI]
10. Andersen MM, Varga S, Folker AP. On the definition of stigma. *Evaluation Clinical Practice*. 2022;28(5):847-853. [DOI]
11. Link BG, Phelan JC. Conceptualizing stigma. *Annu Rev Sociol*. 2001;27:363-385. [DOI]
12. Scambler G. Health-related stigma. *Sociol Heal Illn*. 2009;31(3):441-455. [DOI]
13. Bizumic B, Sheppard H. Prejudice toward people with mental illness. In: Bennett G, Goodall E, editor. *The Palgrave encyclopedia of disability*. Cham: Springer; 2025. p. 1-12. [DOI]
14. Thornicroft G, Sunkel C, Aliev AA, Baker S, Brohan E, El Chammay R, et al. The Lancet Commission on ending stigma and discrimination in mental health. *Lancet*. 2022;400(10361):1438-1480. [DOI][PubMed]
15. Lee YC, Cui Y, Jamieson J, Fu W, Yamashita N. Exploring effects of chatbot-based social contact on reducing mental illness stigma. In: Schmidt A, Väänänen K, Goyal T, Kristensson PO, Peters A, Mueller S, Williamson JR, Wilson ML, editors. *In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*; 2023 Apr 23-28; Hamburg, Germany. New York: Association for Computing Machinery; 2023. p. 1-16. [DOI]
16. Pettigrew TF, Tropp LR. How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *Euro J Social Psych*. 2008;38(6):922-934. [DOI]
17. Sebastian J, Richards D. Changing stigmatizing attitudes to mental health via education and contact with embodied conversational agents. *Comput Hum Behav*. 2017;73:479-488. [DOI]
18. Hanif H, Zulhanis A, Araby S, Rizal Z, Abas K, Othman A, et al. Virtual reality as a bridge to empathy: A case study of the gaza time tunnel VR experience. *Int J Res Innov Soc Sci*. 2025;9(10):6310-6321. [DOI]
19. Obremski D, Akuffo OB, Lücke L, Semine M, Tomiczek S, Weichert HF, et al. Reducing racial bias by interacting with virtual agents: An intervention in virtual reality. *Int Jt Conf Auton Agents Multiagent Syst*. 2023:747-755. [DOI]
20. Zou R, Yin S, Song T, Qin P, Lee YC. Mitigating ageism through virtual reality: Intergenerational collaborative escape room design. In:

- Mueller FF, Kyburz P, Williamson JR, Sas C, editors. In: Extended Abstracts of the CHI Conference on Human Factors in Computing Systems; 2024 May 11-16; Honolulu, USA. New York: Association for Computing Machinery; 2024. p. 1-7. [DOI]
21. Formosa NJ, Morrison BW, Hill G, Stone D. Testing the efficacy of a virtual reality-based simulation in enhancing users' knowledge, attitudes, and empathy relating to psychosis. *Aust J Psychol.* 2018;70(1):57-65. [DOI]
22. Penn DL, Ivory JD, Judge A. The virtual doppelganger: Effects of a virtual reality simulator on perceptions of schizophrenia. *J Nerv Ment Dis.* 2010;198(6):437-443. [DOI]
23. Yuen ASY, Mak WWS. The effects of immersive virtual reality in reducing public stigma of mental illness in the university population of Hong Kong: Randomized controlled trial. *J Med Internet Res.* 2021;23(7):e23683. [DOI]
24. Tay JL, Xie H, Sim K. Effectiveness of augmented and virtual reality-based interventions in improving knowledge, attitudes, empathy and stigma regarding people with mental illnesses-a scoping review. *J Pers Med.* 2023;13(1):112. [DOI][PubMed][PMC]
25. Szekely R, Mason O, Frohlich D, Barley E. The use of virtual reality to reduce mental health stigma among healthcare and non-healthcare students: A systematic review. *Behav Inf Technol.* 2025;44(10):2116-2133. [DOI]
26. Duffy BR, Rooney C, O'Hare GMP, O'Donoghue R. What is a social robot? In: 10th Irish Conference on Artificial Intelligence & Cognitive Science; 1999 Sep 1-3; Cork, Ireland. Dublin: UCD Library; 1999. Available from: <https://researchrepository.ucd.ie/server/api/core/bitstreams>
27. Breazeal C, Dautenhahn K, Kanda T. Social robotics. In: Siciliano B, Khatib O, editors. Springer handbook of robotics. Cham: Springer; 2016. p. 1935-1972. [DOI]
28. Lugin B. Introduction to socially interactive agents. In: Lugin B, Pelachaud C, Traum D, editors. The handbook on socially interactive agents. New York: Association for Computing Machinery; 2021. p. 1-20. [DOI]
29. Ahtinen A, Kaipainen K. Learning and teaching experiences with a persuasive social robot in primary school-findings and implications from a 4-month field study. In: Gram-Hansen SB, Jonassen TS, Midden C, editors. In: 15th International Conference on Persuasive Technology; 2020 Apr 20-23; Aalborg, Denmark. Cham: Springer; 2020. p. 73-84. [DOI]
30. De Carolis BN, Palestra G, Oranger E. Social robots vs. Chatbots: Evaluating the effect as a persuasive technology for children in the healthy eating domain. In: Conati C, Torre I, Volpe G, editors. Proceedings of the 2024 International Conference on Advanced Visual Interfaces; 2024 Jun 3-7; Arenzano, Italy. New York: Association for Computing Machinery; 2024. p. 1-3. [DOI]
31. Donnermann M, Lugin B. Introducing a model for (long-term) personalization of the behavior of a social robot tutor based on self-determination theory and empirical findings. In: 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN); 2024 Aug 26-30; Pasadena, USA. Piscataway: IEEE; 2024. p. 1279-1286. [DOI]
32. De Carolis B, D'Errico F, Macchiarulo N, Rossano V. Investigating the social robots' role in improving children attitudes toward recycling. The case of PeppeRecycle. In: 2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom); 2019 Oct 23-25; Naples, Italy. Piscataway: IEEE; 2019. p. 301-306. [DOI]
33. Ragno L, Borboni A, Vannetti F, Amici C, Cusano N. Application of social robots in healthcare: Review on characteristics, requirements, technical solutions. *Sensors.* 2023;23(15):6820. [DOI][PubMed][PMC]
34. Triantafyllidis A, Alexiadis A, Votis K, Tzovaras D. Social robot interventions for child healthcare: A systematic review of the literature. *Comput Meth Programs Biomed Update.* 2023;3:100108. [DOI]
35. Chang CY, Hwang GJ, Chou YL, Xu ZY, Jen HJ. Effects of robot-assisted digital storytelling on hospitalized children's communication during the COVID-19 pandemic. *Educ Technol Res Dev.* 2023;1-13. [DOI][PubMed][PMC]
36. Rasouli S, Gupta G, Nilsen E, Dautenhahn K. Potential applications of social robots in robot-assisted interventions for social anxiety. *Int J Soc Robot.* 2022;14(5):1-32. [DOI]
37. Castellano G, De Carolis B, D'Errico F, Macchiarulo N, Rossano V. PeppeRecycle: Improving children's attitude toward recycling by playing with a social robot. *Int J Soc Robot.* 2021;13(1):97-111. [DOI]
38. Lo SY, Lai YY, Liu JC, Yeh SL. Robots and sustainability: Robots as persuaders to promote recycling. *Int J Soc Robot.* 2022;14(5):1261-1272. [DOI]
39. Robinson NL, Connolly J, Hides L, Kavanagh DJ. Social robots as treatment agents: Pilot randomized controlled trial to deliver a behavior change intervention. *Internet Interv.* 2020;21:100320. [DOI][PubMed][PMC]
40. Skewes J, Amodio DM, Seibt J. Social robotics and the modulation of social perception and bias. *Phil Trans R Soc B.* 2019;374(1771):20180037. [DOI]
41. Seibt J, Vestergaard C. Fair proxy communication: Using social robots to modify the mechanisms of implicit social cognition. *Res Ideas Outcomes.* 2018;4:e31827. [DOI]
42. Körner J, Chandra S, Dautenhahn K. Can social robots improve people's attitudes toward individuals who stutter? *J Hum-Robot Interact.* 2025;14(4):1-34. [DOI]
43. Appel M, Lugin B, Kühle M, Heindl C. The emotional robotic storyteller: On the influence of affect congruency on narrative transportation, robot perception, and persuasion. *Comput Hum Behav.* 2021;120:106749. [DOI]
44. Spitale M, Okamoto S, Gupta M, Xi H, Matarić MJ. Socially assistive robots as storytellers that elicit empathy. *J Hum-Robot Interact.* 2022;11(4):1-29. [DOI]
45. Steinhäusser SC, Bediako Akuffo O, Weichert HF, Nieding G, Lugin B. Social robots against bullying-effects of embodiment and interactivity on social story experience and efficiency. In: 18th International Conference on Interactive Digital Storytelling; 2025 Dec 1-5; Saint Julian, Malta. Cham: Springer; 2025. p. 72-94. [DOI]

46. Paiva A, Correia F, Oliveira R, Santos F, Arriaga P. Empathy and prosociality in social agents. In: Lugin B, Pelachaud C, Traum D, editors. *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition*. New York: Association for Computing Machinery; 2021. p. 385-432. [DOI]
47. Janoušková M, Tušková E, Weisssová A, Trančík P, Pasz J, Evans-Lacko S, et al. Can video interventions be used to effectively destigmatize mental illness among young people? A systematic review. *Eur Psychiatry*. 2017;41:1-9. [DOI][PubMed]
48. Gonçalves M, Moleiro C, Cook B. The use of a video to reduce mental health stigma among adolescents. *Adolesc Psychiatry*. 2015;5(3):204-211. [DOI]
49. Gomez R, Szapiro D, Galindo K, Merino L, Brock H, Nakamura K, et al. Exploring affective storytelling with an embodied agent. In: 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN); 2021 Aug 8-12; Vancouver, Canada. Piscataway: IEEE; 2021. p. 1249-1255. [DOI]
50. Steinhäusser SC, Riedmann A, Schaper P, Guthmann E, Pfister J, Schmitt K, et al. Second language learning through storytelling with a social robot—an online case study. In: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); 2022 Aug 29-Sep 1; Napoli, Italy. Piscataway: IEEE; 2022. p. 67-74. [DOI]
51. Steinhäusser SC, Maier S, Lugin B. Integrating colored lights into multimodal robotic storytelling. *Empath Comput*. 2025;1(2):202404. [DOI]
52. Alimardani M, Neve L, Verkaart A. Storytelling robots for training of emotion recognition in children with autism; opinions from experts. In: Stephanidis C, Antona M, Ntoa S, editors. *23rd HCI International Conference*; 2021 Jul 24-29. Cham: Springer; 2021. p. 223-233. [DOI]
53. Steinhäusser SC, Maier S, Lugin B. Multimodal robotic storytelling integrating sound effects and background music. *Comput Hum Behav Artif Hum*. 2026;7:100248. [DOI]
54. Habibi A, Damasio A. Music, feelings, and the human brain. *Psychomusicology Music Mind Brain*. 2014;24(1):92-102. [DOI]
55. Herget AK, Frank A. Empfohlene struktur-gleichung? Einfluss von strukturell passender filmmusik auf emotionen und räumliches Präsenzerleben. *Yearb Music Psychol*. 2023. [DOI]
56. Tabak C, Karabulut H. The impact of music on visual storytelling in media. In: *International academic studies in the field of fine arts*. Istanbul: Artikel Akademi; 2024. Available from: https://www.academia.edu/125185189/The_Impact_of_Music_on_Visual_Storytelling_in_Media_M%C3%BCzi%C4%9Fin_Medyada_G%C3%B6rsel_Hikaye_Anlat%C4%B1m%C4%B1_%C3%9Czerindeki_Etkisi
57. Calvert SL, Gersh TL. The selective use of sound effects and visual inserts for children's television story comprehension. *J Appl Dev Psychol*. 1987;8(4):363-375. [DOI]
58. Ge S, Xuan C, Song R, Zou C, Liu W, Zhou J. From text to sound: A preliminary study on retrieving sound effects to radio stories. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*; 2019 Jul 21-25; Paris, France. New York: Association for Computing Machinery; 2019. p. 865-868. [DOI]
59. Miller DW, Marks LJ. Mental imagery and sound effects in radio commercials. *J Advert*. 1992;21(4):83-93. [DOI]
60. Mott RL. *Sound effects: Radio, television and film*. Jefferson: McFarland & Company, Inc.; 2014. Available from: <https://books.google.com.ag/books?id=3vDpAwAAQBAJ&printsec=copyright#v=onepage&q&f=false>
61. Romero-Sanchiz P, Nogueira-Arjona R, Godoy-Ávila A, Gavino-Lázaro A, Freeston MH. Differences in clinical intrusive thoughts between obsessive-compulsive disorder, generalized anxiety disorder, and hypochondria. *Clin Psychol Psychother*. 2017;24(6):O1464-O1473. [DOI][PubMed]
62. Dovidio JF, Major B, Crocker J. Stigma: Introduction and overview. In: Heatherton TF, Kleck RE, Hebl MR, Hull JG, editors. *The social psychology of stigma*. New York: Guilford Press; 2000. p.1-28. Available from: <https://www.researchgate.net/publication/224012639>
63. Goffman E. *Stigma: Notes on the management of spoiled identity*. New York: Simon & Schuster Inc.; 1990. Available from: <https://cdn.penguin.co.uk/dam-assets/books/9780241548011/9780241548011-sample.pdf>
64. Corrigan PW, Watson AC. Understanding the impact of stigma on people with mental illness. *World Psychiatry*. 2002;1(1):16-20. [PubMed][PMC]
65. Augoustinos M, Reynolds KJ. Prejudice, racism, and social psychology. In: *Understanding prejudice, racism, and social conflict*. 1 Oliver's Yard: SAGE Publications Ltd; 2001. p. 1-23. [DOI]
66. Kenny A, Bizumic B, Griffiths KM. The Prejudice towards People with Mental Illness (PPMI) scale: Structure and validity. *BMC Psychiatry*. 2018;18(1):293. [DOI][PubMed][PMC]
67. Bäckström M, Björklund F. Structural modeling of generalized prejudice: The role of social dominance, authoritarianism, and empathy. *J Individ Differ*. 2007;28(1):10-17. [DOI]
68. Bizumic B, Gunningham B, Christensen BK. Prejudice towards people with mental illness, schizophrenia, and depression among mental health professionals and the general population. *Psychiatry Res*. 2022;317:114817. [DOI][PubMed]
69. Aronson E, Wilson TD, Akert RM. *Sozialpsychologie*. Berlin: Springer; 2014. [DOI]
70. Phelan JC, Link BG, Dovidio JF. Stigma and prejudice: One animal or two? *Soc Sci Med*. 2008;67(3):358-367. [DOI]
71. Crisp RJ, Turner RN. Cognitive adaptation to the experience of social and cultural diversity. *Psychol Bull*. 2011;137(2):242-266. [DOI]
72. Vescio TK, Sechrist GB, Paolucci MP. Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *Euro J Social Psych*. 2003;33(4):455-472. [DOI]
73. Cuff BMP, Brown SJ, Taylor L, Howat DJ. Empathy: A review of the concept. *Emotion Rev*. 2016;8(2):144-153. [DOI]
74. Gasioerek J, Ebesu Hubbard AS. Perspectives on perspective-taking in communication research. *Rev Commun*. 2017;17(2):87-105. [DOI]

75. Mae Boag E, Wilson D. Inside experience: Engagement empathy and prejudice towards prisoners. *J Criminal Psychol.* 2014;4(1):33-43. [DOI]
76. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. *J Personality Soc Psychol.* 2000;79(5):701-721. [DOI]
77. Wyse L, Subramanian S. Foundations of interactive sound design for traditional storytelling. In: *Non-Cochlear Sound: Proceedings of the 38th International Computer Music Conference (ICMC); 2012 Sep 9-14; Ljubljana, Slovenia.* Ann Arbor: Michigan Publishing; 2012. Available from: https://lonce.org/home/Publications/publications/WyseSubramanian_Foundations%20of%20Interactive%20Sound%20Design%20for%20traditional%20storytelling.pdf
78. Benish TM, Bramlett RK. Using social stories to decrease aggression and increase positive peer interactions in normally developing pre-school children. *Educ Psychol Pract.* 2011;27(1):1-17. [DOI]
79. Klapow M, Seeff M, Adam M, Greuel M, Hoffman D, Rogge JR, et al. A short, animated storytelling video to reduce addiction stigma: A pilot randomized controlled trial. *Addict Behav Rep.* 2025;22:100622. [DOI]
80. Amsalem D, Greuel M, Liu S, Martin A, Adam M. Effect of a short, animated storytelling video on transphobia among US parents: Randomized controlled trial. *JMIR Public Health Surveill.* 2025;11:e66496. [DOI]
81. Ugwu NF, Ochiaka RE, Asogwa US, Ekwulu EO, Ibenyenwa NC, Onayinka TS, et al. The efficacy of traditional storytelling as a health education intervention strategy in HIV stigma reduction in nigerian rural communities: A randomized controlled trial. *Ikenga.* 2024;25(3):1-24. [DOI]
82. de Graaf A, Sanders J, Hoeken H. Characteristics of narrative interventions and health effects: A review of the content, form, and context of narratives in health-related narrative persuasion research. *Rev Commun Res.* 2016;4:88-131. [DOI]
83. Slater MD, Rouner D. Entertainment—education and elaboration likelihood: Understanding the processing of narrative persuasion. *Commun Theory.* 2002;12(2):173-191. [DOI]
84. Johnson DR. Transportation into a story increases empathy, prosocial behavior, and perceptual bias toward fearful expressions. *Personality Individ Differ.* 2012;52(2):150-155. [DOI]
85. Keer M, van den Putte B, de Wit J, Neijens P. The effects of integrating instrumental and affective arguments in rhetorical and testimonial health messages. *J Health Commun.* 2013;18(9):1148-1161. [DOI][PubMed]
86. Zhuang J, Guidry A. Does storytelling reduce stigma? A meta-analytic view of narrative persuasion on stigma reduction. *Basic Appl Soc Psychol.* 2022;44(1):25-37. [DOI]
87. Christy KR. I, you, or he: Examining the impact of point of view on narrative persuasion. *Medium Psychol.* 2018;21(4):700-718. [DOI]
88. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns.* 2006;61(2):173-190. [DOI][PubMed]
89. Delp C, Jones J. Communicating information to patients: The use of cartoon illustrations to improve comprehension of instructions. *Acad Emerg Med.* 1996;3(3):264-270. [DOI]
90. Marsh AA, Kozak MN, Ambady N. Accurate identification of fear facial expressions predicts prosocial behavior. *Emotion.* 2007;7(2):239-251. [DOI]
91. Wu H, Wang D, Zhou L. Tunes that move us: The impact of music-induced emotions on prosocial decision-making. *Front Psychol.* 2024;15:1453808. [DOI][PubMed][PMC]
92. Hecht M, Kloß A, Bartsch A. Stopping the stigma. How empathy and reflectiveness can help reduce mental health stigma. *Medium Psychol.* 2022;25(3):367-386. [DOI]
93. Beament J. The biology of music. *Psychol Music.* 1977;5(1):3-18. [DOI]
94. Fluckiger B. Sound effects: Strategies for sound effects in film. In: Harper G, Doughty R, Eisentraut J, editors. *Sound and music in film and visual media.* London: Continuum International Publishing; 2009. p. 151-179. Available from: <https://www.zora.uzh.ch/server/api/core/bitstreams/9563e5a3-480d-423e-9a38-015a325dcfbf/content>
95. Thayer JF, Levenson RW. Effects of music on psychophysiological responses to a stressful film. *Psychomusicology.* 1983;3(1):44-52. [DOI]
96. Finnegan R. *Limba stories and story-telling.* Eugene: Wipf and Stock Publishers; 2018. Available from: <https://www.semanticscholar.org/paper/Limba-stories-and-story-telling-Finnegan>
97. Krug HJ. Musik im Hörspiel. In: Schramm H, editor. *Handbuch musik und medien.* Wiesbaden: Springer; 2019. p. 65-93. [DOI]
98. Reddell T. *The sound of things to come: An audible history of the science fiction film.* Minneapolis: University of Minnesota Press; 2018. Available from: <https://www.jstor.org/stable/10.5749/j.ctv69t6p0>
99. Cohen AJ, MacMillan K, Drew R. The role of music, sound effects & speech on absorption in a film: The congruence-associationist model of media cognition. *Can Acoust.* 2006;34(3):40-41. Available from: https://www.researchgate.net/publication/229045922_The_Role_of_Music_Sound_Effects_Speech_on_Absorption_in_a_Film_The_Congruence-Associationist_Model_of_Media_Cognition
100. Rodero E. See it on a radio story: Sound effects and shots to evoked imagery and attention on audio fiction. *Commun Res.* 2012;39(4):458-479. [DOI]
101. Brooke P. *Communicating through story characters: Radio social drama.* Lanham: University Press of America; 1995. Available from: <https://www.amazon.co.uk/Communicating-Through-Characters-Pamela-Brooke/dp/0819199257>
102. Holman T. *Sound for film and television.* 3rd ed. London: Taylor & Francis; 2010. Available from: https://books.google.com/books/about/Sound_for_Film_and_Television

103. Robinson J. The expression and arousal of emotion in music. *J Aesthet Art Crit.* 1994;52(1):13-22. [DOI]
104. Kuch M, Wöllner C. How music can change evaluations of social situations: Exploring the influence of emotional states and perceived group characteristics. *Jahrb Musik.* 2024;32:e187. [DOI]
105. Herget AK, Schramm H. Music as an instrument of nonverbal communication in fictional films. In: 67th Annual Conference of the International Communication Association; 2017 May 25-29; San Diego, USA. Washington: International Communication Association; 2017. [DOI]
106. Jeong JW, Diwadkar VA, Chugani CD, Sinsoongsud P, Muzik O, Behen ME, et al. Congruence of happy and sad emotion in music and faces modifies cortical audiovisual activation. *NeuroImage.* 2011;54(4):2973-2982. [DOI][PubMed]
107. Boltz MG. Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Percept.* 2001;18(4):427-454. [DOI]
108. Bullerjahn C. Wie haben sie den film gehört? Über filmmusik als bedeutungsträger–Eine empirische untersuchung. *Musikpsychologie Jahrb Dtsch Ges Musikpsychol.* 1993;10:140-158. [DOI]
109. Bullerjahn C, Güldenring M. An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology.* 1994;13:99-118. [DOI]
110. Hoeckner B, Wyatt EW, Decety J, Nusbaum H. Film music influences how viewers relate to movie characters. *Psychol Aesthet Creat Arts.* 2011;5(2):146-153. [DOI]
111. Jakubowski K, Margulis EH, Taruffi L. Music-evoked thoughts. *Music Percept Interdiscip J.* 2024;42(1):3-18. [DOI]
112. Cohen AJ. Congruence-Association Model of music and multimedia: Origin and evolution. In: *The psychology of music in multimedia.* Oxford: Oxford University Press; 2013. p. 17-47. [DOI]
113. Costabile KA, Terman AW. Effects of film music on psychological transportation and narrative persuasion. *Basic Appl Soc Psychol.* 2013;35(3):316-324. [DOI]
114. Kock M, Louven C. The power of sound design in a moving picture: An empirical study with emoTouch for iPad. *EMR.* 2019;13:132-148. [DOI]
115. Hegel F, Muhl C, Wrede B, Hielscher-Fastabend M, Sagerer G. Understanding social robots. In: 2009 Second International Conferences on Advances in Computer-Human Interactions; 2009 Feb 1-7; Cancun, Mexico. Piscataway: IEEE; 2009. p. 169-174. [DOI]
116. Lane HC, Schroeder NL. Pedagogical agents. In: *The handbook on socially interactive agents: 20 years of research on embodied conversational agents, intelligent virtual agents, and social robotics volume 2: Interactivity, platforms, application.* New York: Association for Computing Machinery; 2022. p. 307-330. [DOI]
117. Ghafurian M, Muñoz J, Boger J, Hoey J, Dautenhahn K. Socially interactive agents for supporting aging. In: *The handbook on socially interactive agents.* New York: Association for Computing Machinery; 2022. p. 367-402. [DOI]
118. Salem M, Eyssel F, Rohlfing K, Kopp S, Joubin F. Effects of gesture on the perception of psychological anthropomorphism: A case study with a humanoid robot. In: Mutlu B, Bartneck C, Ham J, Evers V, Kanda T, editors. *Social robotics.* Berlin: Springer; 2011. p. 31-41. [DOI]
119. Striepe H, Lugrin B. There once was a robot storyteller: Measuring the effects of emotion and non-verbal behaviour. In: *Social robotics.* Cham: Springer; 2017. p. 126-136. [DOI]
120. Mathur L, Spitale M, Xi H, Li J, Mataric MJ. Modeling user empathy elicited by a robot storyteller. In: 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII); 2021 Sep 28-Oct 1; Nara, Japan. Piscataway: IEEE; 2021. p. 1-8. [DOI]
121. Steinhäusser SC, Piller R, Lugrin B. Combining emotional gestures, sound effects, and background music for robotic storytelling - effects on storytelling experience, emotion induction, and robot perception. In: *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*; 2024 Mar 11-15; Boulder, USA. New York: Association for Computing Machinery; 2024. p. 687-696. [DOI]
122. Li B, Boccanfuso L, Valencia S, Scassellati B, Shic F. Background music and sound effects in human-robot interaction. Worcester: Northeast Robotics Colloquium (NERC); 2015. Available from: https://www.beibinli.com/docs/posters/NERC2015_poster.pdf
123. Robinson F, Pelikan H, Watanabe K, Damiano L, Bown O, Velonaki M. Introduction to the special issue on sound in human-robot interaction. *J Hum-Robot Interact.* 2023;12(4):1-5. [DOI]
124. Zhang J, Sharkey AJC. Listening to sad music while seeing a happy robot face. In: Mutlu B, Bartneck C, Ham J, Evers V, Kanda T, editors. *Social robotics.* Berlin, Heidelberg: Springer; 2011. p. 173-182. [DOI]
125. Steinhäusser SC, Maier S, Lugrin B. Introducing and evaluating a system for an automatic multimodal robotic storyteller featuring the pepper robot. In: 2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); 2025 Aug 25-29; Eindhoven, Netherlands. Piscataway: IEEE; 2025. p. 1523-1530. [DOI]
126. Wong CJ, Tay YL, Wang R, Wu Y. Human-robot partnership: A study on collaborative storytelling. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI); 2016 Mar 7-10; Christchurch, New Zealand. Piscataway: IEEE; 2016. p. 535-536. [DOI]
127. Kwon YS, Lee J, Lee SS. The impact of background music on film audience's attentional processes: Electroencephalography alpha-rhythm and event-related potential analyses. *Front Psychol.* 2022;13:933497. [DOI][PubMed][PMC]
128. Steinhäusser SC, Lugrin B. Integrating sound effects and background music in Robotic storytelling–A series of online studies across different story genres. *Comput Hum Behav Artif Hum.* 2024;2(2):100085. [DOI]
129. Mohammadi G, Park S, Sagae K, Vinciarelli A, Morency LP. Who is persuasive? The role of perceived personality and communication modality in social multimedia. In: *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*; 2013 Dec 9-13; Sydney, Australia. New York: Association for Computing Machinery; 2013. p. 19-26. [DOI]

130. Ma Z, Yang G. Show me a photo of the character: Exploring the interaction between text and visuals in narrative persuasion. *J Heal Commun.* 2022;27(2):125-133. [DOI]
131. Wang S. The power of images: How multimodal hate speech shapes prejudice and prosocial behavioral intentions. *Soc Media Soc.* 2024;10(4):20563051241292990. [DOI]
132. Bonagura A, Abrams D, Teller J. Diagnostic differential between pedophilic-OCD and pedophilic disorder: An illustration with two vignettes. *Arch Sex Behav.* 2022;51(4):2359-2368. [DOI][PubMed]
133. Homonoff Z, Sciutto MJ. The effects of obsession type and diagnostic label on OCD stigma. *J Obsessive Compuls Relat Disord.* 2019;23:100484. [DOI]
134. Seif MN, Winston SM. Tyrannen in meinem kopf: Zwangsgedanken überwinden-ein selbsthilfeprogramm. Paderborn: Junfermann Verlag; 2018. Available from: <https://www.junfermann.de/titel/tyrannen-in-meinem-kopf/1781>
135. Appel M, Gnams T, Richter T, Green MC. The transportation scale–short form (TS–SF). *Medium Psychol.* 2015;18(2):243-266. [DOI]
136. Shen L. On a scale of state empathy during message processing. *West J Commun.* 2010;74(5):504-524. [DOI]
137. Ellard KK, Farchione TJ, Barlow DH. Relative effectiveness of emotion induction procedures and the role of personal relevance in a clinical sample: A comparison of film, images, and music. *J Psychopathol Behav Assess.* 2012;34(2):232-243. [DOI]
138. Appel M, Richter T. Transportation and need for affect in narrative persuasion: A mediated moderation model. *Medium Psychol.* 2010;13(2):101-135. [DOI]
139. Messingschlager TV, Appel M. Creative artificial intelligence and narrative transportation. *Psychol Aesthet Creat Arts.* 2024;18(5):848-857. [DOI]
140. Shen L. The effectiveness of empathy- versus fear-arousing antismoking PSAs. *Health Commun.* 2011;26(5):404-415. [DOI][PubMed]
141. Rodríguez-de-Dios I, Soto-Sanfiel MT. Reducing transphobia with the narratives of transgender YouTubers. *Cyberpsychology.* 2024;18(5):2. [DOI]
142. Shen X, Pai YS, Kiuchi D, Bao K, Aoki T, Meguro H, et al. Dementia eyes: Co-design and evaluation of a dementia education augmented reality experience for medical workers. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*; 2023 Apr 23-28; Hamburg, Germany. New York: Association for Computing Machinery; 2023. p. 1-18. [DOI]
143. Plutchik R. A psychoevolutionary theory of emotions. *Soc Sci Inf.* 1982;21(4-5):529-553. [DOI]
144. Youssef K, Said S, Alkork S, Beyrouthy T. A survey on recent advances in social robotics. *Robotics.* 2022;11(4):75. [DOI]
145. Allen A, Savery R, Robinson N. Consequential sounds and their effect on human robot interaction. In: *Sound and robotics: Speech, non-verbal audio and robotic musicianship*. Boca Raton: Chapman and Hall/CRC; 2023. p. 101-127. [DOI]
146. Balzotti A, De Carolis B, Massaro S, Perla L, Rossano V. Healthy pepper: Nutritional education through social robotics and storytelling. In: *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*; 2023 Jul 10-13; Orem, USA. Piscataway: IEEE; 2023. p. 121-123. [DOI]
147. De Carolis B, D'Errico F, Rossano V. Pepper as a storyteller: Exploring the effect of human vs. Robot voice on children's emotional experience. In: *Human-computer interaction–INTERACT 2021*. Cham: Springer; 2021. p. 471-480. [DOI]
148. Steinhäusser SC, Zehe A, Maier S, Hotho A, Lugin B. Development of the fully automatic robotic storyteller–comparing manual, semi-automatic, and LLM-based emotional story annotation for multimodal robotic storytelling. *IEEE Trans Affect Comput.* 2026:1-18. [DOI]
149. Siegman AW, Boyle S. Voices of fear and anxiety and sadness and depression: The effects of speech rate and loudness on fear and anxiety and sadness and depression. *J Abnorm Psychol.* 1993;102(3):430-437. [DOI][PubMed]
150. Oh J, Jeong SY, Jeong J. The timing and temporal patterns of eye blinking are dynamically modulated by attention. *Hum Mov Sci.* 2012;31(6):1353-1365. [DOI][PubMed]
151. Stern JA, Walrath LC, Goldstein R. The endogenous eyeblink. *Psychophysiology.* 1984;21(1):22-33. [DOI][PubMed]
152. Holland MK, Tarlow G. Blinking and thinking. *Percept Mot Skills.* 1975;41(2):403-406. [DOI]
153. Drew GC. Variations in reflex blink-rate during visual-motor tasks. *Q J Exp Psychol.* 1951;3(2):73-88. [DOI]
154. Brannick SF, LaCroix AN. Blinking indexes dynamic attending during and after music listening. *Sci Rep.* 2025;15(1):27262. [DOI][PubMed][PMC]
155. JASP. Version 0.95.4 [Software]. 2025. Available from: <https://jasp-stats.org/previous-versions/>
156. Marzi G, Balzano M, Marchiori D. K-Alpha Calculator-Krippendorff's Alpha Calculator: A user-friendly tool for computing Krippendorff's Alpha inter-rater reliability coefficient. *MethodsX.* 2024;12:102545. [DOI][PubMed][PMC]
157. Blanca M, Alarcón R, Arnau J, Bono R, Bendayan R. Non-normal data: Is ANOVA still a valid option? *Psicothema.* 2017;4(29):552-557. [DOI]
158. Finch H. Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated. *Methodology.* 2005;1(1):27-38. [DOI]
159. Santavirta S, Paranko B, Seppälä K, Hyönä J, Nummenmaa L. Modelling human social vision with cinematic stimuli. *BioRxiv [Preprint]*. 2024. [DOI]
160. Andreu-Sánchez C, Martín-Pascual MÁ, Gruart A, Delgado-García JM. Viewers change eye-blink rate by predicting narrative content. *Brain Sci.* 2021;11(4):422. [DOI]
161. Steinhäusser SC, Heckel M, Lugin B. The way you see me - comparing results from online video-taped and in-person robotic storytelling research. In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*; 2024 Mar 11-15; Boulder, USA. New York: Association for Computing Machinery; 2024. p. 1018-1022. [DOI]

162. Cummings JJ, Tsay-Vogel M, Cahill TJ, Zhang L. Effects of immersive storytelling on affective, cognitive, and associative empathy: The mediating role of presence. *New Medium Soc.* 2022;24(9):2003-2026. [DOI]
163. McDonald B, Böckler A, Kanske P. Soundtrack to the social world: Emotional music enhances empathy, compassion, and prosocial decisions but not theory of mind. *Emotion.* 2022;22(1):19-29. [DOI]
164. Ochanda DA. State of mental health of students in universities. In: Aloka P, editor. *Mental health crisis in higher education*. Hershey: IGI Global Scientific Publishing; 2023. p. 14-28. [DOI]
165. Sheppard H, Bizumic B, Caele A. Prejudice toward people with borderline personality disorder: Application of the prejudice toward people with mental illness framework. *Int J Soc Psychiatry.* 2023;69(5):1213-1222. [DOI][PubMed][PMC]
166. Na JJ, Park JL, LKhagva T, Mikami AY. The efficacy of interventions on cognitive, behavioral, and affective public stigma around mental illness: A systematic meta-analytic review. *Stigma Health.* 2022;7(2):127-141. [DOI]
167. Birtel MD, Oldfield G. Affective, cognitive, and behavioral mental illness stigma in health care: A comparison between general ward nurses and the general population. *Stigma Health.* 2022;7(4):380-388. [DOI]
168. Vaughn LA, Hesse SJ, Petkova Z, Trudeau L. "This story is right on": The impact of regulatory fit on narrative engagement and persuasion. *Euro J Social Psych.* 2009;39(3):447-456. [DOI]
169. Green MC, Jenkins KM. Interactive narratives: Processes and outcomes in user-directed stories. *J Commun.* 2014;64(3):479-500. [DOI]
170. Kennedy-Hendricks A, McGinty EE, Summers A, Krenn S, Fingerhood MI, Barry CL. Effect of exposure to visual campaigns and narrative vignettes on addiction stigma among health care professionals: A randomized clinical trial. *JAMA Netw Open.* 2022;5(2):e2146971. [DOI][PubMed][PMC]
171. Krendl AC, Pescosolido BA. Countries and cultural differences in the stigma of mental illness: The east-west divide. *J Cross Cult Psychol.* 2020;51(2):149-167. [DOI]
172. Ociskova M, Prasko J, Cerna M, Jelenova D, Kamaradova D, Latalova K, et al. Obsessive compulsive disorder and stigmatization. *Act Nerv Super Rediv.* 2013;55(1-2):19-26. Available from: <https://rediviva.sav.sk/55i1/19.pdf>
173. Coles ARL, Lahey CA, Fawcett EJ, Fawcett JM. Overcoming the forbidden: Identification and stigma of unacceptable thoughts in obsessive-compulsive disorder. *Stigma Health.* 2023:314-325. [DOI]

Publisher's Note

Science Exploration remains a neutral stance on jurisdictional claims in published maps and institutional affiliations. The views expressed in this article are solely those of the author(s) and do not reflect the opinions of the Editors or the publisher.