Facilitating Episodic Simulation in Anxiety: Role of Sensory Scaffolding and Scenario Modality

Julie L. Ji¹ · M. Joseph Meyer² · Bethany A. Teachman²

Published online: 26 March 2020
© The Author(s) 2020

Abstract

Cognitive bias modification of interpretation style (CBM-I) is a family of cognitive training programs that seek to reduce anxious thinking by training people to assign relatively more positive meanings to ambiguous situations. CBM-I's effects may be enhanced by encouraging more vivid imagery-based episodic simulation of events and by increasing engagement with the training materials. This study investigated the role of *sensory scaffolding* (whether pictures, or pictures + sound were added) and verbal *scenario modality* (whether scenarios were delivered visually or aurally) on episodic simulation (Vivid; Plausible; Changing Perspective ratings) and user engagement (Relatable, Comprehensible, Enjoyable ratings). Amazon Mechanical Turk workers ($N = 187$) with varied anxiety symptom severity *read* or *listened* to brief scenarios that varied by sensory scaffolding and verbal scenario modality. Results were somewhat mixed. Generally, picture scaffolding tended to facilitate both episodic simulation and user engagement (relative to no scaffolding), irrespective of scenario modality and anxiety level.

Keywords Episodic simulation · User engagement · Anxiety · Cognitive bias modification of interpretation

Anxiety is associated with elevated expectancies of negative outcomes for ambiguous situations, despite the rare occurrence of such outcomes (MacLeod and Byrne 1996; Miranda et al. 2008; Miranda and Mennin 2007). Such expectancies can contribute to emotional distress and behavioral avoidance in individuals with anxiety disorders

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41811-020-00070-x>) contains supplementary material, which is available to authorized users.

✉ Julie L. Ji
julie.ji@uwa.edu.au

¹ School of Psychological Science, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia

² Department of Psychology, University of Virginia, Charlottesville, VA, USA

(Boswell et al. 2013; McEvoy and Erceg-Hurn 2015). Cognitive interventions developed to tackle such maladaptive expectancies typically do so via therapist-driven cognitive restructuring (identifying and re-evaluating unhelpful thoughts) in therapy. Recently, technology-delivered cognitive training approaches, such as cognitive bias modification of interpretation style (CBM-I; Mathews and Mackintosh 2000), have been developed to directly modify the interpretation of uncertain situations in more positive and flexible ways.

A widely used CBM-I paradigm is the ambiguous situations task (Mathews and Mackintosh 2000), in which participants practice repeatedly resolving verbal scenarios depicting ambiguous everyday situations in a benign or positive manner. Meta-analytic evidence suggests that CBM-I experimental and treatment designs that instruct participants to imagine themselves in the situations described in each scenario (episodic simulation) (Szpunar et al. 2014) may be particularly promising for modifying bias and reducing anxiety symptoms (Menne-Lothmann et al. 2014). However, results on the effectiveness of CBM-I treatment studies for reducing anxiety and depression symptoms have been mixed (Menne-Lothmann et al. 2014). One reason for the mixed results may be that anxious individuals struggle with episodic simulation of positive, but not negative, events (Wu et al. 2015). Therefore, enhancing positive episodic simulation may increase intervention efficacy.

The present study investigated task-related factors that could facilitate episodic simulation and user-engagement processes in the context of materials used in CBM-I interventions for anxiety. Episodic simulation processes refer to the extent the imagined situations depicted in CBM-I scenarios were experienced as vivid, plausible, and resulted in perspective change. User-engagement processes refer to the extent the imagined situations depicted in CBM-I scenarios were relatable and comprehensible, and the extent the task of imagining the scenarios was enjoyable.

Promoting Healthy Episodic Simulation to Reduce Anxiety

Episodic simulation refers to mental imagery-based constructions of specific hypothetical experiences (Schacter et al. 2008), based on the flexible recombination of information stored in memory (Irish and Piguet 2013; Schacter and Addis 2007). Mental imagery refers to the experience of perception in the absence of external sensory input (Kosslyn et al. 2001), and its neural bases overlap significantly with that of actual sensory perception (Kosslyn 1994; Pearson and Kosslyn 2015). Therefore, episodic simulation can construct the perceptual features of hypothetical experiences and evoke the experiential correlates of such experiences (Lang 1979; Moulton and Kosslyn 2009). Episodic simulation affords individuals the capacity to *predict* and *pre-experience* future events (Addis et al. 2007; Schacter et al. 2007), thus playing a key role in judgment, decision-making, emotion-regulation, and goal-directed behavior (Bulley et al. 2016; Moulton and Kosslyn 2009; Seligman et al. 2013).

In healthy individuals, training the ability to simulate specific alternative outcomes vividly can help to weaken the plausibility of initially negative outcome expectancies (Jing et al. 2017). One way to enhance the vividness of episodic simulation is through repeated simulation (Szpunar and Schacter 2013). However, for individuals with generalized anxiety disorder (GAD), repeated simulation does not appear to enhance

the vividness (ease of generation and scene detail) of positive future events, but does increase the vividness of negative events (Wu et al. 2015). As such, alternative approaches to simple repetition may be required to facilitate the promotion of positive episodic simulation processes in the treatment of anxiety.

Enhancing Episodic Simulation in Cognitive Bias Modification

One promising approach that has been developed, CBM-I, involves practicing new and non-threatening ways to interpret ambiguous anxiety-provoking situations. More than simply repetitively imagining positive events, CBM-I scenarios are specifically designed such that the outcome of each situation remains ambiguous until the final word. This ambiguity feature aims to evoke anxiety-linked maladaptive negative outcome expectancies, which are subsequently refuted, promoting the learning of healthier, less rigid cognitive and emotional responding to uncertainty. Recent meta-analytic evaluation of CBM-I interventions has shown that CBM-I interventions that instruct participants to imagine themselves in the scenarios were significantly more effective than CBM-I interventions without such instruction (Menne-Lothmann et al. 2014). However, CBM treatment studies are not always effective, especially at reducing anxiety symptoms (Cristea et al. 2015), and there are problems with engagement as participants report the intervention to be boring (Beard et al. 2012). Therefore, understanding ways to enhance healthy episodic simulation and user engagement may enhance the effectiveness of CBM-I treatments.

Effects of Sensory Scaffolding

Within the context of CBM-I research, episodic simulation is typically evoked using verbal sentences that describe a situation as it unfolds over time, enabling experimental control over how and when the ambiguity in the situation is resolved. Typically, the final word serves to resolve the situation positively or negatively (Mathews and Mackintosh 2000). To successfully engage with such interventions, participants must comprehend the verbal sentences while engaging in simultaneous scene construction, maintenance, updating, and visualization (Hassabis et al. 2007). This complex process may be made easier if the background scene was provided for participants to construct the rest of the scenes by reducing cognitive burden.

The use of sensory scaffolds in the form of background pictures has been used in some CBM-I treatment studies targeting depression. For example, Yiend et al. (2014) reported deciding to provide background pictures related to the scenario topic following feedback from pilot participants that the pictures helped them to imagine themselves in the situation. In Blackwell et al. (2015), background picture scaffolding was provided in conjunction with word phrases, and participants were asked to combine the two cues into a plausible imagined situation. Further, an online CBM-I treatment study for anxiety in an adolescent sample compared the effectiveness and user experience of *picture-word* training scenarios against the standard CBM-I *sentence descriptor* training scenarios (de Voogd et al. 2017). Unexpectedly, *picture-word* training scenarios were rated as less enjoyable, harder to concentrate on, and more confusing than

sentence descriptor training scenarios (de Voogd et al. 2017). However, it is difficult to interpret these results in relation to CBM-I training targeting ambiguity, as *picture-word* stimuli do not involve ambiguity resolution given the emotional outcome is denoted by the word phrase. Therefore, *picture-word* training's impact on anxiety likely operates via different mechanisms than CBM-I *sentence descriptor*-based training. Also, *picture-word* cues may be particularly high in cognitive load given this training does not provide descriptions of an unfolding event, instead requiring participants to generate such episodic details themselves, which may partly explain why this training is harder to follow than CBM-I *sentence descriptor* training.

Importantly, no study has examined multi-sensory scaffolds, such as background pictures *with* background sounds. Given that anxious samples that may be more readily overwhelmed by task demands (see Steinman and Teachman 2015), a systematic evaluation of the role of sensory scaffolding in episodic simulation of stimuli used in CBM-I targeting ambiguity is required to inform optimal approaches for modifying interpretation style in anxiety and increasing episodic simulation more generally.

Effects of Scenario Modality

Another factor that may influence episodic simulation processes is cue modality. Though some studies have presented scenarios aurally (Hayes et al. 2010; Murphy et al. 2007), the majority of CBM-I training use visually presented scenarios (see Menne-Lothmann et al. 2014). Working memory research suggests that visual mental imagery vividness is reduced by concurrent tasks that tax visuospatial working memory (Baddeley and Andrade 2000). Thus, it is possible that compared with hearing the scenarios, reading the scenarios may interfere with one's ability to simultaneously imagine the unfolding situation, particularly for anxious individuals. On the other hand, listening requires the verbal information to be comprehended in real time, which may be more challenging than reading the same information because individuals can dwell on parts of the text that require additional processing (Ferreira and Anes 1994; Lund 1991).

Few studies have directly compared visual and auditory presentation of training scenarios. In one laboratory CBM-I training study, stimuli modality did not affect the magnitude of change in interpretation style or emotional reactivity to a stressor (Standage et al. 2010). However, participants in the auditory presentation condition, but not the visual presentation condition, reported an increase in state depressed mood, attributed to longer testing times required for auditory presentation, relative to self-paced visual presentation, of scenarios. Further investigations are required to understand how scenario presentation modality impacts CBM-I training efficacy and effectiveness. An important area to examine is how scenario modality may impact episodic simulation and user engagement-related, particularly when sensory scaffolds are in use.

Anxiety Symptom Level

While previous research suggests anxiety symptoms may be negatively related to simulation vividness as well as likelihood judgments and interpretations of positive outcomes (Wu et al. 2015), it is unclear whether anxiety symptoms will also be related

to user engagement with the training materials. By sampling across individuals with varying levels of anxiety symptoms, the present study also aimed to investigate how anxiety severity moderates the effects of sensory scaffold and scenario modality. Including participants with the full range of anxiety symptoms is also consistent with recommendations from the National Institute of Mental Health's Research Domain Criteria (RDoC; NOT-MH-11-005 NIMH Research Domain Criteria (RDoC): Interim Guidance) framework to assess the impact of heterogeneity in symptom severity.

The current study

Given that CBM-I training is designed to induce learning via expectancy violation through repeated exposure to scenarios depicting benign endings to potentially anxiety-provoking situations, greater experiential engagement with the scenarios may enhance CBM-I training efficacy. The aim of the current study was to assess which method of scenario presentation is associated with the highest level of episodic simulation of and user engagement with the scenarios. As such, we tested three levels of sensory scaffolding (none, visual only, visual + auditory), which represented increasing levels of experiential cueing. In addition, for each level of sensory scaffolding, we tested whether scenario modality (reading vs. listening) moderated any effects of sensory scaffolding. Our aim was to assess which format of presenting the scenario was associated with optimal effects on our indicators of episodic simulation and user engagement, based on our belief that greater episodic simulation and user engagement will result in stronger CBM-I training effects and greater adherence (this latter connection needs to be tested in a future study, but our goal here was to first see if we could improve episodic simulation and user engagement, given we view them as viable mechanisms to enhance learning in interventions like CBM-I).

To assess the strength and impact of episodic simulation, we assessed reported: (a) vividness of the imagined scenario (Vivid); (b) plausibility judgment of the scenario's positive ending (Plausible); and (c) perspective change resulting from the scenario's ending (Changing Perspective¹). To assess user engagement with the training materials, we assessed: (d) scenario relatability (Relatable); (e) scenario ease of comprehension (Comprehensible); and (f) how fun/enjoyable reading/hearing the different types of scenarios (while imagining them) were (Enjoyable).

Hypotheses

It was hypothesized that sensory scaffolding would facilitate episodic simulation of potentially anxiety-provoking situations that end positively. Specifically, it was predicted that greater sensory scaffolding would be associated with higher Vivid, Plausible, and Changing Perspective ratings, even after the effect of anxiety symptom level was taken into account. Similarly, greater sensory scaffolding was expected to be associated with greater user engagement (higher Relatable,

¹ The label "Changing Perspective" is used for expedience and reflects a subjective evaluation of the extent a perspective change occurred, as opposed to an objective evaluation of perspective shift.

Comprehensible, Enjoyable ratings) even after the effect of anxiety symptom level was taken into account. For scenario modality, given that both reading and listening have advantages and disadvantages, the study made no directional hypotheses regarding its main or interaction effects with scenario scaffolding.

Finally, the study explored whether anxiety symptom severity moderated the effects of episodic scaffolding and scenario modality on all outcome variables. Regarding the zero-order relationships between anxiety symptom severity and outcome variables based on prior research, it was predicted that higher anxiety symptoms would be related to lower Vividness and Plausible ratings, but higher Changing Perspective ratings due to the likelihood that those with a higher (vs. lower) anxiety symptom level would presumably find the positive scenario endings more surprising. Tests of the relationships between anxiety symptom level and user engagement outcomes were exploratory.

Using a 3 (Sensory Scaffold) $\times$ 2 (Scenario Modality) experimental design among participants with variable levels of anxiety symptoms, the present study aimed to illuminate task-related factors that influence general, as well as anxiety-specific, episodic simulation and user engagement processes, which may inform translational research on promoting healthier episodic simulation among anxious samples.

Method

Participants

A sample of $N=197$ adults were recruited over the Internet through Amazon's Mechanical Turk (MTurk). Demographics information is presented in Table 1. Anxiety symptom level was assessed via the Depression, Anxiety, Stress Scales-Short Form Anxiety Subscale (DASS-21 AS; Lovibond and Lovibond 1995). DASS-21 AS scores were then multiplied by 2 to enable comparison with normative data on the 42-item full version of the DASS anxiety subscale (DASS-AS). To assess participants across the full spectrum of anxiety symptom levels, the study aimed to recruit equivalent numbers of individuals with moderate to high anxiety (equivalent to a score of ≥ 10 on the DASS-AS) and individuals with low anxiety (equivalent to a score of < 10 on the DASS-AS; 10 was used as the cutoff following Lovibond and Lovibond 1995). MTurk workers who completed the anxiety screener and had a HIT Approval Rate² of $> 95\%$ were allowed to participate in the study. No other eligibility criteria were enforced. The task took approximately 30-min to complete, and participants were compensated \$4.00 for completing the task. The study was approved by the University of Virginia Institutional Review Board for Social and Behavioral Sciences (protocol number 2015-0125-00).

² The Human Intelligence Task (HIT, referring to a task put on MTurk, such as the current study) approval rate is akin to a quality rating for a worker. The Approval Rate is the rate of HITs (or tasks) that Requesters (experimenters) have approved for each Worker, indicating satisfactory completion of a HIT.

Table 1 Participant characteristics

Variable	<i>M (SD)</i>	Range/%
DASS Anxiety score	10.96 (10.12)	0–42
Age (years)	38.01 (11.81)	20–71
Gender (Female)	128	0.69
Education		
Some high school	1	0.54%
High School graduate	21	11.29%
Some college	74	39.78%
Undergraduate degree	64	34.41%
Graduate degree	26	13.98%
Ethnicity		
American Indian/Alaskan Native	2	1.07%
Black/African origin	15	8.06%
East Asian	4	2.15%
South Asian	4	2.15%
White	158	84.95%
Other/unknown	3	1.61%
Race		
Hispanic	10	5.38%
Non-Hispanic	170	91.40%
Other/unknown	6	3.23%
Language		
Native	171	91.94%
Fluent	14	7.53%
Proficient	1	0.54%

DASS-AS score = Depression Anxiety and Stress Scale – Anxiety subscale score

Materials

Episodic Simulation Stimuli

Twenty-four previously rated scenarios used in CBM-I training studies in the lab were used as stimuli to evoke episodic simulation (see Appendix 1 for the full list of scenarios). The scenarios depict everyday situations that are potentially anxiety provoking, but all end positively or benignly. Half of the scenarios comprised social threat-relevant content (e.g., giving a presentation; dinner with colleagues), and half comprised physical-threat-relevant content (e.g., waiting for the results of an annual physical; feeling panic-related sensations during a flight). Equal numbers of social and physical-threat scenarios were allocated to each experimental condition.

Half of the scenarios were visually presented (Scenario Modality *Read* condition), and half were aurally presented (Scenario Modality *Listen* condition), read out by a female native American-English speaker. Within each Scenario Modality condition, a third of scenarios were presented without any sensory scaffold (No Scaffold), a third

were presented with a preceding background picture related to the scene (Picture Scaffold), and a third were presented with a preceding background picture *and* background sound related to the scene (Picture + Sound Scaffold). Sensory scaffolding materials were background pictures and background sounds created by undergraduate research assistants at the University of Virginia or taken from a Google Image search (with no use restrictions). All participants received the same set of scenarios, which were randomly selected from a large pool of scenarios used in previous CBM-I intervention studies and assigned to a modality condition. To increase comparability across conditions, we assigned scenarios to condition based on the criteria that each condition had to have the same number of scenario type, i.e., those occurring in social vs. non-social situations, involving physical vs. social threat.

Episodic Simulation Task

Each scenario was three sentences long, and the final sentence ended with a word fragment (a word with one letter missing). The outcome of all scenarios was ambiguous until the final word fragment was completed, which disambiguated the scenarios in a non-threatening way. For example, participants might read, “As you are walking down a crowded street, you see your neighbor on the other side. You call out, but she does not answer you. Standing there in the street, you think that this must be because she was distract_ed.” Participants would type the letter “t” to complete the word “distracted.” The word fragment was presented in the same way for aurally and visually presented scenarios. In both conditions, the final word appeared on the screen (with one letter missing). Thus, for auditory scenarios, participants saw nothing on the screen while they were listening to the scenario, until the final word, which popped up on the screen visually. Following each scenario, participants answered the following questions: (a) “How vividly did you imagine the scenario (as if you were really there and experiencing it first-hand)?” (Vivid); (b) “How easy was it to follow the story?” (Comprehensible); (c) “To what extent did this story’s ending make you see this situation in a new way?” (Changing Perspective); (d) “To what extent did this scenario’s ending feel possible, like it could really happen?” (Plausible); and (e) “To what extent did you feel you could relate to the situations that were presented?” (Relatable). Response ratings were provided on a five-point Likert scale ranging from 1 (not at all) to 5 (totally). After completion of all trials, participants completed user engagement ratings for each type of scenario they encountered in response to the question, “How fun/enjoyable was X the story” (Enjoyable), where X denoted the type of scenario it was (e.g., seeing a picture + listening to the story).³

Scenarios were presented in six blocks representing the six experimental conditions. Participants provided all ratings at the end of each scenario trial, except for Enjoyable ratings, which were provided at the end of the task in relation to each type of scenarios participants encountered (e.g., “scenarios presented visually with a picture but no sound”). Block order was randomized across participants.

³ Scenario type 1: X = *Reading* the story only; 2: X = Seeing a picture + *reading* the story; 3: X = Seeing a picture + hearing background noise + *reading* the story; 4: X = *Listening* to the story only; 5: X = Seeing a picture + *listening* to the story; 6: X = Seeing a picture + hearing background noise + *listening* to the story

Anxiety Measure

Anxiety symptom level was assessed via a DASS-21 AS (Lovibond and Lovibond 1995) screener before the main study. In the DASS-21 AS, participants are asked how often they “felt” seven anxiety symptoms (relating to apprehension, panic, and worry) in the past week, on a scale of 0 (“did not apply to me at all”) to 3 (“applied to me very much, or most of the time”). The DASS-21 has strong psychometric properties, including high reliability and concurrent validity (Antony et al. 1998).

Procedure

This study was conducted over the Internet via Amazon’s MTurk. Participants were informed that: “This study looks at how reading and listening to stories online affect how the story is experienced. It takes about 25-30 minutes. In the task, you will be presented with stories about everyday situations and asked to imagine yourself experiencing them. After each story, you will be asked to rate your experience.” Interested participants completed an initial consent form and the DASS-21 AS screener, for which they were not compensated. If eligible and interested in doing the main study, participants then completed a second consent form and demographics information via a survey on Qualtrics, which concluded with a link to complete the main task.

In the study, participants were provided with the following instructions on the computer screen: “In this task, you will read or listen to a series of short stories. Please imagine yourself in the situation described in each story. We will show you a demonstration of this with a fun “Lemon” exercise, coming up next. Remember, even if the story describes you reacting in a way that you would not usually react, please try to picture yourself responding in the way the story describes. There will be an incomplete word at the end of each paragraph. Press the key on the keyboard that completes the word. When you correctly complete the word, you will move on to the next screen and be asked a series of questions about the story.” Participants then completed a guided imagery exercise involving interactions with a lemon in their hands, and before starting the study, the following instructions were displayed: “In this study, you will see a set of short stories. All stories will start with the story’s name, so you know what the story is about. For some stories, you will READ or LISTEN to the story straight after the story’s name is shown. For other stories, you will SEE a background PICTURE related to the story BEFORE you READ or LISTEN to the story. For the remaining stories, you will SEE a background PICTURE and HEAR background NOISE related to the story BEFORE you READ or LISTEN to the story. Before we start, please make sure your sound volume is turned on.”

Following completion of the main task, participants were presented with debriefing information and paid for their participation.

Data Analysis Plan

Linear mixed effects models were used to examine the effects of Sensory Scaffolding (none; picture; picture + sound), Scenario Modality (visual; auditory) and DASS-AS score (continuous) on each of the dependent variables (Vivid; Plausible; Changing

Perspective; Relatable; Comprehensible; Enjoyable). Analyses were conducted using the “lme4” package in R (RStudio Team 2018; Bates et al. 2014), with Sensory Scaffolding, Scenario Modality, and DASS-AS as fixed effects, and Subject as random intercept. Two effect sizes, as represented by R^2 , were computed for each model, one for fixed effects only (marginal) and one for fixed plus random effects (conditional). R^2 were computed in R using the approach recommended by Nakagawa and Schielzeth (2012). Significant three-way interactions were further decomposed into two-way interactions for each Scenario Modality condition separately. Significant two-way interactions were further decomposed into simple effects with Tukey’s HSD pairwise multiple comparison corrections using the “multcomp” package (Hothorn et al. 2017) and “emmeans” package (Lenth et al. 2018) in R. In addition, DASS-AS scores were centered at 10 (denoting moderate anxiety) to evaluate effects at an established cutoff point for classifying severity (Lovibond and Lovibond 1995).

Results

Data Reduction

Of the 197 participants, ten participants were found to have completed the task more than once. These ten participants were excluded and no data from them were included in the analyses. Of the remaining 187 participants, one did not complete demographics information, and two did not complete the DASS screener, but their data were retained for analyses.

Effects of Sensory Scaffolding, Scenario Modality, and Anxiety Symptoms on Episodic Simulation

Table 2 provides descriptive statistics for the six outcome variables by scenario type.

Summary Table of Results

Table 3 summarizes results for the main and interaction effects of Sensory Scaffolding, Scenario Modality, and their interactions with Anxiety Symptoms on all outcome variables.

Vivid

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 0.43\%$ for fixed effects, and $R^2_{\text{conditional}} = 71.97\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality

Analyses revealed a non-significant trend towards a main effect of Sensory Scaffolding, $\chi^2(2) = 5.92$, $p = 0.051$. Tukey’s HSD pairwise comparisons revealed significantly lower Vivid ratings in the No Scaffold condition compared with the Picture Scaffold

Table 2 Descriptive statistics of the six outcome variables by Scenario Modality and Sensory Scaffold conditions

Scenario modality	Sensory scaffold	Episodic simulation outcomes						User engagement outcomes					
		Vivid		Plausible		Changing Perspective		Relatable		Comprehensible		Enjoyable	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Visual	No Scaffold	4.05	0.80	4.05	0.75	2.64	1.02	3.90	0.82	4.49	0.59	2.92	1.14
	Picture Scaffold	4.16	0.68	4.00	0.75	2.78	1.00	4.05	0.74	4.52	0.61	3.67	1.01
	Picture and Sound Scaffold	4.12	0.72	3.94	0.81	2.83	1.05	3.97	0.80	4.50	0.59	4.09	1.00
Auditory	No Scaffold	4.11	0.74	3.77	0.81	3.03	1.02	3.82	0.83	4.48	0.60	3.36	1.07
	Picture Scaffold	4.18	0.73	3.71	0.89	3.00	0.94	3.86	0.81	4.42	0.67	4.05	0.93
	Picture and Sound Scaffold	4.08	0.72	4.02	0.78	2.81	1.05	3.98	0.81	4.49	0.64	4.50	0.86

Table 3 Summary of results for the main and interaction effects of Sensory Scaffolding and Scenario Modality, and their interactions with Anxiety Symptom level on Episodic Simulation and User Engagement outcome variables

Conditions	Episodic simulation processes			User engagement		
	Vivid	Plausible	Changing Perspective	Relatable	Comprehensible	Enjoyable
Sensory Scaffolding	Pic > None	NS	Pic > Pic + Sound > None	Pic and Pic + Sound > None	NS	Pic + Sound > Pic > None
Scenario Modality	NS	Read > Listen	Listen > Read	NS	NS	Listen > Read
Sensory Scaffolding x Scenario Modality	NS	Read: Pic + Sound = Pic = None Listen: Pic + Sound > Pic and None	Read: Pic + Sound and Pic > None Listen: Pic and None > Pic + Sound	Read: Pic > None Listen: Pic + Sound > Pic and None	NS	NS
Moderation by Anxiety Symptom Level	Higher Anxiety = Higher Vividness: Listen > Read	NS	NS	NS	Higher anxiety = Lower Ease: Pic < None	NS

Abbreviated terms: None, No Scaffold; Pic, Background Picture Scaffold; Pic + Sound, Background Picture and Background Sound Scaffold; Read, Visually presented scenarios; Listen, Aurally presented scenarios; NS, not statistically significant at $p < .05$

condition ($b = -0.07$, $p = 0.03$, 95%C.I.[-0.13: -0.02]), but not compared with Vivid ratings for the Picture + Sound Scaffold condition ($b = -0.01$, $p = 0.91$, 95%C.I.[-0.06: 0.04]). Vivid ratings for the Picture and Picture + Sound Scaffold conditions did not differ significantly ($b = 0.06$, $p = 0.07$, 95%C.I.[0.01: 0.12]). There was no significant main effect of Scenario Modality, $\chi^2(1) = 1.63$, $p = 0.20$, nor any interaction effects between Sensory Scaffolding and Scenario Modality, $\chi^2(2) = 2.34$, $p = 0.31$, as depicted in Fig. 1.

Moderation by Anxiety Symptoms

A two-way interaction between Anxiety Symptoms and Stimuli Modality was found, $\chi^2(1) = 3.86$, $p = 0.049$, reflecting a greater positive relationship between DASS-AS score and Vivid ratings for auditory relative to visual scenarios. However, decomposition of this interaction indicated that the relationships between DASS-AS score and Vivid ratings were not significantly different from zero for either Visual scenarios ($b = -0.001$, C.I.: -0.01: 0.01) or Auditory scenarios ($b = 0.002$, C.I.: -0.01: 0.01). No other effects were statistically significant, all $\chi^2 < 2.63$, all $p > 0.20$.

Summary

Picture scaffolding was superior to No Scaffolding for scenario vividness, irrespective of scenario modality. In addition, *listening* compared with *reading* scenarios was associated with a relatively greater positive relationship between anxiety level and scenario vividness, but the effect is small and there was no significant relationship between anxiety symptom level and vividness in either modality condition alone.

Fig. 1 Mean Vivid ratings (in response to the question “How vividly did you imagine the scenario (as if you were really there and experiencing it first-hand)”) as a function of Sensory Scaffolding and Scenario Modality. Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

Plausible

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 3.85\%$ for fixed effects, and $R^2_{\text{conditional}} = 58.78\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality

Analyses revealed a main effect of Scenario Modality, $\chi^2(1) = 29.64$, $p < 0.001$, reflecting higher Plausible ratings for Visual scenarios relative to Auditory scenarios ($b = 0.17$, $p < 0.001$, 95% C.I. [0.11: 0.23]). This main effect was subsumed within a 2-way interaction between Sensory Scaffolding and Scenario Modality, $\chi^2(2) = 33.38$, $p < 0.001$, shown in Fig. 2. Tukey's HSD pairwise comparisons revealed that Plausible ratings did not differ significantly across Sensory Scaffolding level for Visual scenarios (all $b \leq |0.12|$, all $p \geq 0.24$). In contrast, for Auditory scenarios, compared with the Picture + Sound Scaffold condition, Plausible ratings were significantly lower in the No Scaffold condition, $b = -0.25$, $p < 0.001$, 95% C.I. [-0.12: 0.09], and the Picture Scaffold condition, $b = -0.32$, $p < 0.001$, 95% C.I. [-0.36: -0.14]. Overall, Plausible ratings in the Auditory No Scaffold condition and Auditory Picture Scaffold condition were significantly lower than for all other types of scenarios, all $b \geq |0.17|$, all $p \leq 0.02$, although there was no significant difference between the Visual vs. Auditory Picture + Sound conditions, $b = -0.07$, $p = 0.07$, 95% C.I. [-0.43: -0.22].

Fig. 2 Mean Plausible ratings (in response to the question “To what extent did this scenario’s ending feel possible, like it could really happen?”) as a function of Sensory Scaffolding and Scenario Modality. Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

Moderation by Anxiety Symptoms

DASS-AS score was not involved in any main or interaction effects, all $\chi^2 \leq 2.34$, all $p \geq 0.12$.

Summary

Listening to scenarios was associated with lower scenario *plausibility* relative to reading scenarios, except when there was maximal scaffolding (picture + sound). These effects were independent of anxiety symptom level.

Changing Perspective

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 2.63\%$ for fixed effects, and $R^2_{\text{conditional}} = 68.19\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality

Analyses revealed a main effect of Sensory Scaffolding, $\chi^2(2) = 10.04$, $p = 0.006$, such that Changing Perspective ratings were highest for the Picture Scaffold condition, followed by the Picture + Sound Scaffold condition, with the lowest being for the No Scaffold condition, although no follow-up tests for differences between Sensory Scaffolding conditions were statistically significant, all $b \leq |0.07|$, all $p \geq 0.22$. A main effect of Scenario Modality was also found, $\chi^2(1) = 43.22$, $p < 0.001$, in which

Fig. 3 Mean Changing Perspective ratings (in response to the question “To what extent did this scenario’s ending make you see this situation in a new way?”) as a function of Sensory Scaffolding and Scenario Modality. Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

Changing Perspective ratings for Visual scenarios was significantly lower than for Auditory scenarios, $b = -0.19$, $p < 0.001$, 95%C.I.[-0.26: -0.13].

These two main effects were further subsumed under a two-way interaction, $\chi^2(2) = 24.14$, $p < 0.001$, depicted in Fig. 3. Tukey's HSD pairwise comparisons revealed that, for Visual scenarios, Changing Perspective ratings in the No Scaffold condition were significantly lower than the Picture + Sound Scaffold condition, $b = -0.18$, $p = 0.03$, 95%C.I.[-0.30: -0.07], but not lower than the Picture Scaffold condition, $b = -0.13$, $p = 0.22$, 95%C.I.[-0.25: -0.02]. There was no significant difference between the Picture Scaffold and Picture + Sound Scaffold conditions, $b = -0.05$, $p = 0.96$, 95%C.I.[-0.17: 0.07]. In contrast, for Auditory scenarios, compared with the Picture + Sound Scaffold condition, Changing Perspective ratings were significantly higher for the Picture + Sound Scaffold condition relative to both the No Scaffold condition, $b = 0.23$, $p = 0.002$, 95%C.I.[0.11: 0.35], and the Picture Scaffold condition, $b = 0.19$, $p = 0.02$, 95%C.I.[0.07: 0.31], with no significant differences between No Scaffold and Picture Scaffold condition, $b = 0.04$, $p = 0.98$, 95%C.I.[-0.08: 0.16]. Overall, Changing Perspective ratings in the Auditory No Scaffold and Auditory Picture Scaffold conditions were significantly higher than for all other conditions (though the difference between Visual Picture + Sound vs. Auditory Picture scaffold scenarios did not reach statistical significance, $b = -0.16$, $p = 0.06$, 95%C.I.[-2.76: 0.06], all other $b \geq |0.19|$, $p \leq 0.02$). There was no difference between Visual vs. Auditory Picture + Sound Scaffold scenarios, $b = 0.03$, $p = 0.99$.

Moderation by Anxiety Symptoms

DASS-AS scores were not reliable in any main or interaction effects, all $\chi^2 \leq 2.45$, all $p \geq 0.12$.

Summary

Maximal scaffolding was associated with the lowest perspective change when scenarios were aurally presented, but highest when scenarios were visually presented. Aural presentation was superior to visual presentation in general, except in the maximal scaffolding condition. These effects were independent of anxiety symptom level.

Effects of Sensory Scaffolding, Scenario Modality, and Anxiety Symptoms on User Engagement

Relatable

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 1.69\%$ for fixed effects, and $R^2_{\text{conditional}} = 67.87\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality Analyses revealed a main effect of Sensory Scaffolding, $\chi^2(2) = 10.10$, $p = 0.006$, where Relatable ratings were significantly lower in the No Scaffold condition than the Picture Scaffold condition, $b = -0.09$, $p = 0.03$, 95%C.I.[-0.15: -0.02], and the Picture + Sound Scaffold

Fig. 4 Mean Relatable ratings (in response to the question “To what extent did you feel you could relate to the situation that was presented?”) as a function of Sensory Scaffolding and Scenario Modality. Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

condition, $b = -0.11$, $p = 0.003$, 95% C.I. $[-0.17; -0.04]$), but there were no significant differences between the Picture Scaffold and Picture + Sound Scaffold conditions, $b = -0.02$, $p = 0.80$. This main effect was subsumed under a two-way interaction between Sensory Scaffolding $\times$ Scenario Modality, $\chi^2(2) = 11.32$, $p = 0.003$, as depicted in Fig. 4.

Tukey's HSD pairwise comparisons revealed that, for Visual scenarios, Relatable ratings for the No Scaffold condition were significantly lower than the Picture Scaffold condition, $b = -0.15$, $p = 0.02$, 95% C.I. $[-0.24; -0.06]$, but not the Picture + Sound Scaffold condition, $b = -0.05$, $p = 0.81$. The Picture vs. Picture + Sound Scaffold conditions did not differ significantly, $b = 0.09$, $p = 0.38$. For Auditory scenarios, Relatable ratings in the No Scaffold condition were significantly lower than the Picture + Sound Scaffold condition, $b = -0.15$, $p = 0.01$, 95% C.I. $[-0.24; -0.06]$. Ratings for the Picture Scaffold condition were also lower than the Picture + Sound Scaffold condition, $b = -0.13$, $p = 0.05$, 95% C.I. $[-0.22; -0.04]$. The No Scaffold and Picture Scaffold conditions did not differ, $b = -0.02$, $p = 0.99$, 95% C.I. $[-0.11; 0.07]$. Overall, Relatable ratings in the Visual Picture Scaffold condition were significantly higher than the Visual and Auditory No Scaffold conditions, and the Auditory Picture Scaffold condition, all $b \geq |0.14|$, all $p \leq 0.02$. There was no difference between the Visual vs. Auditory Picture + Sound conditions, $b = -0.01$, $p = 0.99$.

Moderation by Anxiety Symptoms

DASS-AS scores were not reliable in any main or interaction effects, all $\chi^2 \leq 2.63$, all $p \geq 0.10$.

Summary

Relatability was highest with picture scaffolding if scenarios were visually presented, and highest with maximal scaffolding if scenarios were aurally presented. These effects were independent of anxiety symptom level.

Comprehensible

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 1.23\%$ for fixed effects, and $R^2_{\text{conditional}} = 74.73\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality

Sensory Scaffolding and Scenario Modality were not involved in any main or interaction effects, all $\chi^2 \leq 4.19$, all $p \geq 0.12$, as depicted in Fig. 5.

Moderation by Anxiety Symptoms

Analyses yielded a main effect of DASS-AS score, $\chi^2(1) = 4.79$, $p = 0.029$, such that higher DASS-AS scores were associated with *lower* Comprehensible ratings. This main effect of DASS-AS was subsumed within a two-way interaction with Sensory Scaffolding, $\chi^2(2) = 7.99$, $p = 0.018$. Tukey's HSD pairwise comparisons showed that the negative relationship between DASS-AS score and Comprehensible ratings was significantly smaller in the Picture Scaffold condition than the No Scaffold condition

Fig. 5 Mean Comprehensible ratings (in response to the question “How easy was it to follow the story?”) as a function of Sensory Scaffolding and Scenario Modality (X = scenario type; e.g., seeing a picture + reading the story). Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

($b = -0.05$, $p = 0.048$). No other pairwise comparisons were statistically significant (all $b \leq |0.003|$, all $p \geq 0.22$), and no slope differed from 0 (95% C.I. [-0.01: 0.002]).

Summary

Ease of comprehension did not differ as a function of sensory scaffolding level or scenario modality. However, higher anxiety symptom level was associated with lower ease of comprehension, although this negative effect of anxiety was buffered by picture scaffolding relative to no scaffolding.

Enjoyable

For the overall generalized linear mixed effects model, $R^2_{\text{marginal}} = 21.25\%$ for fixed effects, and $R^2_{\text{conditional}} = 46.79\%$ for the full model.

Effects of Sensory Scaffolding and Scenario Modality Analyses revealed a main effect of Sensory Scaffold, $\chi^2(2) = 191.80$, $p < 0.001$, in which Enjoyable ratings in the Picture Scaffold and Picture + Sound Scaffold conditions were significantly higher than the No Scaffold condition, and Enjoyable ratings in the Picture + Sound Scaffold condition were also significantly higher than in the Picture Scaffold condition (all $b = 0.43$, all $t > 7.12$, all $p < 0.001$). Also, a main effect of Scenario Modality was found, such that Enjoyable ratings were significantly lower for Visual scenarios than for

Fig. 6 Mean Enjoyable ratings (in response to the question “How fun/enjoyable was X the stories?”) as a function of Sensory Scaffolding and Scenario Modality (X = scenario type; e.g., seeing a picture + reading the story). Error bars denote standard errors. Ratings ranged from 1 (Not at all) to 5 (Totally)

Auditory scenarios, $b = -0.41$, $p < 0.001$, 95% C.I. $[-0.50; -0.31]$. Results are depicted in Fig. 6.

Moderation by Anxiety Symptoms

DASS-AS scores were not involved in any main or interaction effects, all $\chi^2 \leq 2.68$, all $p \geq 0.26$.

Summary

Scenarios with greater sensory scaffolding were more fun/enjoyable than those with less scaffolding, and listening was more fun/enjoyable than reading scenarios. Overall, listening to scenarios with maximal scaffolding was the most fun/enjoyable, and reading scenarios without scaffolding was the least fun/enjoyable.

Relationship between Anxiety Symptoms and Episodic Simulation and User Engagement Outcomes

Spearman's correlations between DASS-AS scores and all outcome variables are presented in Table 4. Contrary to hypotheses, higher anxiety was not related to lower Vivid ratings. Consistent with hypotheses, higher anxiety was associated with lower Plausible and higher Changing Perspective ratings. Higher anxiety was also associated with lower Relatable, and lower Comprehensible ratings, but anxiety symptom level was not related to Enjoyable ratings.

Discussion

In the context of verbal scenarios used in CBM-I training, the present study presents preliminary findings on the role of sensory scaffolding and stimuli modality in facilitating episodic simulation and user engagement among individuals with variable levels of anxiety symptoms. Although results differed across the outcome measures, overall results indicated that scaffolding (relative to no scaffolding) was associated with higher *vividness*, *plausibility*, and *perspective change*, and higher *relatability* and *enjoyableness*, providing general support for the sensory scaffolding hypothesis. In addition, while the effects of sensory scaffold were moderated by scenario modality in complex ways, reading or listening to scenarios was inconsequential to scenario *plausibility*, *relatability*, *perspective change*, and *enjoyableness* when scenarios were preceded by *maximal scaffolding* (background picture and sound). However, when scenarios were preceded by *picture scaffold* only, *listening* to scenarios was associated with *lower* scenario ending *plausibility* and scenario *relatability*, but higher *perspective change* and *enjoyableness*, relative to *reading* scenarios.

Interestingly, anxiety symptom level did not correlate with vividness ratings in the present study, which appears to contradict previous findings showing anxiety to be associated with the reduced vividness of imagined positive events (Wu et al. 2015). However, previous research did not involve the use of sensory scaffolds. In the present

Table 4 Correlations between DASS-AS Score and Episodic Simulation and User Engagement outcome variables

	DASS Anxiety Score		Vivid	Plausible	Changing Perspective	Relatable	Comprehensible	Enjoyable
DASS Anxiety Score	Spearman's rho	—	−0.003	−0.087	0.079	−0.076	−0.091	−0.026
	<i>p</i> value	—	0.912	0.004	0.009	0.012	0.003	0.382
Vivid	Spearman's rho		—	0.463	0.238	0.605	0.598	0.173
	<i>p</i> value		—	<.001	<.001	<.001	<.001	<.001
Plausible	Spearman's rho			—	0.170	0.752	0.440	0.208
	<i>p</i> value			—	<.001	<.001	<.001	<.001
Changing Perspective	Spearman's rho				—	0.228	0.071	0.182
	<i>p</i> value				—	<.001	0.018	<.001
Relatable	Spearman's rho					—	0.482	0.255
	<i>p</i> value					—	<.001	<.001
Comprehensible	Spearman's rho						—	0.151
	<i>p</i> value						—	<.001
Enjoyable	Spearman's rho							—
	<i>p</i> value							—

*p < .05; **p < .01; ***p < .001

study, *listening* compared with *reading* scenarios was associated with a relatively greater positive relationship between anxiety level and scenario vividness, but the effect is small and there was no significant relationship between anxiety symptom level and vividness when examining either modality condition alone. It is possible that having been exposed to scenarios that included pictorial and auditory scaffolds, participants learned to more vividly imagine scenarios in general, thereby reducing overall individual differences in scenario imagery vividness. Future research can disentangle this possible effect by presenting non-scaffolded trials before scaffolded trials and assessing anxiety-linked individual differences in imagery vividness for the non-scaffolded trials and scaffolded trials separately.

Finally, while higher anxiety symptoms were associated with lower *ease of comprehension* of scenarios, this negative effect was reduced for scenarios with *picture scaffold*, but not *picture + sound scaffold*, relative to *no scaffold*, indicating a beneficial effect of *picture scaffolding* on episodic simulation for anxious populations.

Implications for Episodic Simulation Research

The present results provide preliminary evidence that providing relevant background pictures and sounds may facilitate episodic simulation of potentially anxiety-provoking everyday situations in a non-threatening way, irrespective of their anxiety level. Specifically, the provision of sensory information may help individuals simulate such events more vividly, increase the subjective plausibility of the non-threatening endings (particularly if scenarios were read from the computer screen), and result in greater likelihood of seeing the situation in a new way (particularly if scenarios were listened to via headphones).

Importantly, the present results suggest that while scenarios with *maximal scaffolding* (both background picture and sound) were the most *enjoyable*, it was not consistently associated with better outcomes in other domains. For example, *maximal scaffolding* was worse than no scaffolding for *plausibility* (when scenarios were read), and for *perspective change* (when scenarios were listened to). In contrast, irrespective of scenario modality, *picture-only scaffolding* was consistently associated with either similar outcomes or significantly better outcomes compared with *no scaffolding*. Therefore, recognizing that factors that increase the enjoyableness of episodic simulation do not necessarily align with factors that facilitate episodic simulation, *picture-only scaffolding* may confer benefits to episodic simulation and user engagement with few downsides. Alternatively, rather than modifying the scenarios in a way that works optimally on average, it may be helpful to examine the utility of matching individual learning/scaffolding preferences with the preferred sensory scaffolding level and scenario modality.

It should be noted that the results are preliminary and the mixed nature of findings suggests there may be no obvious one-size-fits-all approach to improving CBM-I training. Decisions for training design may require prioritizing improving some features at the possible expense of others.

Limitations

Several factors constrain the interpretation of results from the present study. First, given the within-subject nature of the study design, observed differences in outcome variables

likely reflect relative differences, rather than absolute differences, between scenario types. As such, results may be especially relevant for informing the design of multisession CBM-I intervention studies, where scenario conditions can be delivered in order of increasing *relative* fun/enjoyment level to incentivize continued engagement with the intervention in the same individuals over time. In addition, although the nature of anxiety-related threat (social and physical) depicted in scenarios was equivalent across conditions and scenarios were previously rated for equivalence in valence, the scenarios were presented in a fixed format. The decision to use different specific scenarios (with matched threat content and valence) across conditions was considered critical to avoid repetition of materials, which would have contaminated ratings, especially for items like plausibility and perspective change. Also, the choice to use a within-subject design was prioritized to allow comparisons across conditions for the fun/enjoyment outcome. That said, future research could replicate and extend the present study by using a between-subjects design and randomizing scenarios to condition.

In addition, it must be acknowledged that the outcome variables of interest were based on subjective beliefs about the outcomes, rather than objective indicators (e.g., participants reflected on their belief that the training materials were comprehensible vs. demonstrating actual comprehension). For variables that require metacognitive insight, such as the degree of perspective change, scenario comprehensibility, and level of engagement with scenario content, inclusion of more objective assessments would be valuable. For example, memory assessments of training content at the end of each training session may be informative. Analogously, to determine how the subjective ratings of enjoyment align with more objective measures, it may be helpful to assess how enjoyment ratings predict persistence in CBM-I training (e.g., number of scenarios or training sessions completed). Note, while correct answers to the missing letters during training (and to the comprehension questions that are typically also presented) might also be helpful to evaluate, we suspect these metrics would have limited utility to demonstrate depth of engagement for most scenario-based CBM-I paradigms given rates for correct responses tend to be high, resulting in ceiling effects.

Another potential limitation is that the study was conducted using an online crowdsourcing platform (Mechanical Turk), which involved non-probability sampling. However, direct comparison studies have shown MTurk samples are more representative of the national population than college student samples and community samples recruited from college towns (Berinsky et al. 2012). Further, many studies have shown MTurk data to be high in both reliability (Behrend et al. 2011; Buhrmester et al. 2011; Johnson and Borden 2012) and validity (Behrend et al. 2011; Mason and Suri 2012), similar to those from large representative samples (Shapiro et al. 2013) and student populations (Johnson and Borden 2012; Wickham et al. 2015). Also, the online nature of MTurk studies appears to facilitate willingness to disclose psychologically relevant and sensitive information (Andover 2014; Shapiro et al. 2013; Wurtele et al. 2013).

An additional limitation of conducting the study over the internet is that participants undertook the study in non-controlled settings unknown to the experimenter. The study did not include additional catch items or comprehension questions to assess continued engagement in part because previous research suggest that the employment of attention checks do not necessarily guarantee increased attention, but may have drawbacks such as increased attrition (Berinsky et al. 2016) and interference on natural cognitive responses on tasks (Hauser and Schwarz 2016). Instead, we followed recommended

guidelines (Buhrmester et al. 2018) and restricted participation to those with a 95% approval rate or higher, which has been demonstrated to be equally effective at reducing inattention as the use of attention checks (Peer et al. 2014). Notably, performance (percentage correct) for the word fragment completion component of the task was above 97%, indicating adequate attention to task. Finally, given CBM-I interventions are often delivered over the internet, it was appropriate that the present study was also delivered over the internet.

Given the current study was experimental (rather than a test of the efficacy of an episodic simulation-based CBM-I intervention itself), it remains to be seen if sensory scaffolding can improve intervention efficacy (e.g., cognitive bias change and anxiety symptom reduction) and user engagement and retention. Also, baseline interpretation bias level was not measured, so it was not possible to evaluate whether the scenario manipulations had different effects as a function of variation in participants' levels of interpretation bias. Finally, background pictures used in this study were not tailored to each participant, which may have limited the extent to which training scenarios could plausibly reflect each participant's natural environment. The use of personalized images in future studies may enhance the effects of sensory scaffolding.

More generally, interpretation of the study's effect sizes should be treated with caution given the lack of consensus about their calculation and interpretation for generalized linear mixed effects models (Nakagawa and Schielzeth 2012). It should be noted that while conditional R^2 values (for fixed and random effects) for the overall models were generally high, marginal R^2 values (for fixed effects only) were low for all outcome variables except for *enjoyableness*. This indicates substantial variability between participants in how the various scenario presentation formats were received. Thus, it will be important for future research to determine what the optimal scenario presentation format is for each individual. Variability in the scenario presentation effects was also evident across the outcome measures; specifically, the scenario presentation effects explained the most variance in *enjoyableness* of the scenarios, followed by *plausibility* and *perspective change*.

Future Directions

We assume that CBM-I's effectiveness rests partly on participants' capacity to vividly and realistically pre-experience the situations depicted in the training scenarios. This assumption draws from meta-analytic findings showing that studies with instructions for participants to vividly imagine themselves experiencing the training scenarios is associated with enhanced CBM-I effectiveness (Menne-Lothmann et al. 2014). This vivid experiential representation is expected to activate negative expectancies and, in turn, maximize learning when such expectancies are refuted at the end of the scenario. From a clinical perspective, we expect that the extent to which anxious participants can vividly simulate the scenarios, find the situations plausible, and experience perspective changes will enhance CBM-I outcomes, but this link remains to be tested. Evaluating the extent to which the six outcome variables in the current study relates to CBM-I success is of critical theoretical and applied importance, and represents a priority for future research.

The present study represents a first step in understanding how scenario presentation-related factors may impact episodic simulation and user engagement-related outcomes.

An important next step would be to investigate how the various scenario presentation conditions differ in relation to the six outcome variables using a between-group design in the context of actual CBM-I training, which will also serve to illuminate how variation in the six outcome variables relate to CBM-I efficacy and effectiveness (e.g., changes in interpretation bias and anxiety symptoms). That is, future research should test whether enhancing episodic simulation via the provision of sensory scaffolding can enhance CBM-I treatment effects by making it easier to imagine training scenarios in a vivid, plausible, and relatable way, and in turn enhance the efficacy of CBM-I. The current study lays the groundwork to help determine the optimal method to use.

Based on CBM-I attrition data and feedback from participants in our studies, we know that user-engagement processes (such as ease of comprehension and enjoyment of training) are also important for intervention success. Further evaluations are required to ascertain the relative importance of these outcome features in improving outcomes of CBM-I interventions. Also, future studies investigating the impact of sensory scaffolding and scenario modality factors should use clinically anxious samples to determine the clinical implications of these outcomes. Future translational research could evaluate the effect of sensory scaffolding on episodic simulation in populations that experience impairments in episodic simulation, such as individuals with depression (Gamble et al. 2019) or dementia (Irish and Piolino 2016).

While this study sought to understand if the effects of scaffolding and modality were the same across different levels of anxiety, it may be fruitful for future studies to examine other individual differences moderators, such as episodic simulation ability, or depression. We also note that highly anxious individuals may find benign, more neutral scenario outcomes more plausible and relatable than overtly positive scenario outcomes (Alden et al. 2004; Murphy et al. 2007). Therefore, future research testing the effects of sensory scaffolding in CBM-I interventions for anxiety should also evaluate the magnitude of such effects across different levels of scenario positivity to identify the optimal training scenario parameters.

In addition, although an exciting outcome for the present line of research is potentially improving the success of CBM-I interventions for anxiety, we believe the CBM-I approach has applications beyond reducing anxiety in high anxious populations. For instance, in the case of anxiety, CBM-I designs can be used to understand the optimal level of anxiety required for motivating adaptive precautionary action, such as applying sunscreen (Notebaert et al. 2014) and bush fire preparation (Notebaert et al. 2017). Ultimately, understanding how to enhance the impact of script-evoked episodic simulation is of potential relevance to any area of research that engages episodic simulation processes.

Finally, an important aspect of the present study is the timing of sensory scaffolding delivery. A key aspect of emotion-evoking mental imagery is that the individual is able to self-generate and elaborate on the sensory-experiential details of the imagined stimulus (Lang 2016). The aim of sensory scaffolding in this study was to help individuals begin this process, rather than to eliminate the need to generate such sensory-experiential details themselves. We believe this is important, as one purpose of imagery-based CBM-I is to help individuals develop more helpful modes of anticipating upcoming anxiety-provoking situations, which occurs before they encounter the actual situation (i.e., involving mental

representation of, rather than in-vivo exposure to, the feared situation). Thus, to maximize the opportunity to self-generate sensory-experiential details, visual and/or auditory scaffolds were provided before scenario delivery, rather than throughout scenario delivery. Future research should investigate the potential trade-offs between sensory scaffolding and self-generation of imagined sensory-experiential situational details during episodic simulation.

Conclusion

Although results were mixed, the present research suggests that the provision of background pictures as sensory scaffolding can reliably facilitate episodic simulation as well as user engagement (as compared with no scaffolding), irrespective of scenario modality and anxiety level. While aural presentation of scenarios plus sensory scaffolding seems advisable overall, the optimal level of scaffolding may depend on the outcome variable of interest. Understanding stimuli-related factors that impact episodic simulation processes is important for both basic and translational research in order to strengthen the effects of our interventions, especially those delivered via technology.

Acknowledgments J. L. Ji was supported by the Forrest Research Foundation. B. A. Teachman was supported by R34MH106770 and R01MH113752. We are thankful to Emily Holmes, Simon Blackwell, Bundy Mackintosh, Andrew Mathews, and their research teams for sharing some of their training and testing materials. Some of their scenarios were included, typically in a modified form, as part of our materials. We are thankful to Daniel H. Funk and Diheng Zhang for their technical assistance in conducting this study.

Funding Information This work was supported in part by NIMH grants (R34MH106770 and R01MH113752) to B. Teachman. B. Teachman has a significant financial interest in Project Implicit, Inc., which provided services in support of this project under contract with the University of Virginia.

Compliance with Ethical Standards

Conflict of Interest B. A. Teachman has a significant financial interest in Project Implicit, Inc., which provided services in support of this project under contract with the University of Virginia.

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

References

- Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, 45(7), 1363–1377. <https://doi.org/10.1016/j.neuropsychologia.2006.10.016>.

- Alden, L. E., Mellings, T. M., & Laposa, J. M. (2004). Framing social information and generalized social phobia. *Behaviour Research and Therapy*, 42(5), 585–600. [https://doi.org/10.1016/S0005-7967\(03\)00163-3](https://doi.org/10.1016/S0005-7967(03)00163-3).
- Andover, M. S. (2014). Non-suicidal self-injury disorder in a community sample of adults. *Psychiatry Research*, 219(2), 305–310. <https://doi.org/10.1016/j.psychres.2014.06.001>.
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, 10(2), 176.
- Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*. US: American Psychological Association. <https://doi.org/10.1037/0096-3445.129.1.126>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4*. <https://doi.org/10.18637/jss.v067.i01>.
- Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients' attitudes toward cognitive bias modification (CBM): a qualitative study. *Behavioural and Cognitive Psychotherapy*, 40(5), 618–633. <https://doi.org/10.1017/S1352465811000671>.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43(3), 800. <https://doi.org/10.3758/s13428-011-0081-0>.
- Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3), 351–368. <https://doi.org/10.1093/pan/mpr057>.
- Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2016). Can we turn shirkers into workers? *Journal of Experimental Social Psychology*, 66, 20–28. <https://doi.org/10.1016/j.jesp.2015.09.010>.
- Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., et al. (2015). Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: a randomized controlled trial. *Clinical Psychological Science*, 3(1), 91–111. <https://doi.org/10.1177/2167702614560746>.
- Boswell, J. F., Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2013). Intolerance of uncertainty: a common factor in the treatment of emotional disorders. *Journal of Clinical Psychology*, 69(6), 630–645. <https://doi.org/10.1002/jclp.21965>.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5. <https://doi.org/10.1177/1745691610393980>.
- Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological Science*, 13(2), 149–154. <https://doi.org/10.1177/1745691617706516>.
- Bulley, A., Henry, J., & Suddendorf, T. (2016). Prospection and the present moment: the role of episodic foresight in intertemporal choices between immediate and delayed rewards. *Review of General Psychology*, 20(1), 29–47. <https://doi.org/10.1037/gpr0000061>.
- Cristea, I., Kok, R. N., & Cuijpers, P. (2015). The efficacy of cognitive bias modification interventions for mental health problems: a meta-analysis. *European Psychiatry*, 30, 850. [https://doi.org/10.1016/S0924-9338\(15\)30663-5](https://doi.org/10.1016/S0924-9338(15)30663-5).
- de Voogd, E. L., de Hullu, E., Burnett Heyes, S., Blackwell, S. E., Wiers, R. W., & Salemink, E. (2017). Imagine the bright side of life: a randomized controlled trial of two types of interpretation bias modification procedure targeting adolescent anxiety and depression. *PLoS One*, 12(7), e0181147. <https://doi.org/10.1371/journal.pone.0181147>.
- Ferreira, F., & Anes, M. (1994). Why study spoken language? In *Handbook of psycholinguistics* (pp. 33–56). San Diego: Academic Press.
- Gamble, B., Moreau, D., Tippet, L. J., & Addis, D. R. (2019). Specificity of future thinking in depression: a meta-analysis. *Perspectives on Psychological Science*, 1745691619851784. <https://doi.org/10.1177/1745691619851784>.
- Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *The Journal of Neuroscience*, 27(52), 14365–14374. <https://doi.org/10.1523/JNEUROSCI.4549-07.2007>.
- Hauser, D. J., & Schwarz, N. (2016). Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 48(1), 400–407. <https://doi.org/10.3758/s13428-015-0578-z>.

- Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy*, 48(3), 171–178. <https://doi.org/10.1016/j.brat.2009.10.006>.
- Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., Scheibe, S., & Hothorn, M. T. (2017). Package ‘multcomp’. *CRAN*.
- Irish, M., & Piguet, O. (2013). The pivotal role of semantic memory in remembering the past and imagining the future. *Frontiers in Behavioral Neuroscience*, 7, 27. <https://doi.org/10.3389/fnbeh.2013.00027>.
- Irish, M., & Piolino, P. (2016). Impaired capacity for prospection in the dementias – theoretical and clinical implications. *The British Journal of Clinical Psychology*, 55(1), 49–68. <https://doi.org/10.1111/bjc.12090>.
- Jing, H. G., Madore, K. P., & Schacter, D. L. (2017). Preparing for what might happen: an episodic specificity induction impacts the generation of alternative future events. *Cognition*, 169, 118–128. <https://doi.org/10.1016/j.cognition.2017.08.010>.
- Johnson, D. R., & Borden, L. A. (2012). Participants at your fingertips: using Amazon’s Mechanical Turk to increase student–faculty collaborative research. *Teaching of Psychology*, 39(4), 245–251. <https://doi.org/10.1177/0098628312456615>.
- Kosslyn, S. M. (1994). *Image and brain: the resolution of the imagery debate*. Cambridge: MIT Press.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews. Neuroscience*, 2, 635. <https://doi.org/10.1038/35090055>.
- Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16(6), 495–512. <https://doi.org/10.1111/j.1469-8986.1979.tb01511.x>.
- Lang, P. J. (2016). Imagery in therapy: an information processing analysis of fear–republished article. *Behavior Therapy*, 47(5), 688–701. <https://doi.org/10.1016/j.beth.2016.08.011>.
- Lenth, R., Singmann, H., Love, J., Buurkner, P., & Herve, M. (2018). Emmeans: estimated marginal means, aka least-squares means. 2018. *R package version*, 1(4).
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U).
- Lund, J. R. (1991). A comparison of second language listening and reading comprehension. *The Modern Language Journal*, 75(2), 196–204. <https://doi.org/10.1111/j.1540-4781.1991.tb05350.x>.
- MacLeod, A. K., & Byrne, A. (1996). Anxiety, depression, and the anticipation of future positive and negative experiences. *Journal of Abnormal Psychology*, 105(2), 286. <https://doi.org/10.1037/0021-843X.105.2.286>.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behavior Research Methods*, 44(1), 1–23. <https://doi.org/10.3758/s13428-011-0124-6>.
- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology*, 109(4), 602–615. <https://doi.org/10.1037/0021-843X.109.4.602>.
- McEvoy, P. M., & Erceg-Hum, D. M. (2015). The search for universal transdiagnostic and trans-therapy change processes: evidence for intolerance of uncertainty. *Journal of Anxiety Disorders*, 41, 96–107. <https://doi.org/10.1016/j.janxdis.2016.02.002>.
- Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kavanova, Z., Haller, S. P., Drukker, M., et al. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS One*, 9(6). <https://doi.org/10.1371/journal.pone.0100925>.
- Miranda, R., & Mennin, D. S. (2007). Depression, generalized anxiety disorder, and certainty in pessimistic predictions about the future. *Cognitive Therapy and Research*, 31(1), 71–82. <https://doi.org/10.1007/s10608-006-9063-4>.
- Miranda, R., Fontes, M., & Marroquin, B. (2008). Cognitive content-specificity in future expectancies: role of hopelessness and intolerance of uncertainty in depression and GAD symptoms. *Behaviour Research and Therapy*, 46(10), 1151–1159. <https://doi.org/10.1016/j.brat.2008.05.009>.
- Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 364(1521), 1273–1280. <https://doi.org/10.1093/acprof:oso/9780195395518.003.0040>.
- Murphy, R., Hirsch, C. R., Mathews, A., Smith, K., & Clark, D. M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behaviour Research and Therapy*, 45(7), 1517–1529. <https://doi.org/10.1016/j.brat.2007.01.007>.
- Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>.

- Notebaert, L., Chrystal, J., Clarke, P. J., Holmes, E. A., & MacLeod, C. (2014). When we should worry more: using cognitive bias modification to drive adaptive health behaviour. *PLoS One*, 9(1), e85092. <https://doi.org/10.1371/journal.pone.0085092>.
- Notebaert, L., Tilbrook, M., Clarke, P. J., & MacLeod, C. (2017). When a bad bias can be good: anxiety-linked attentional bias to threat in contexts where dangers can be avoided. *Clinical Psychological Science*, 5(3), 485–496. <https://doi.org/10.1177/2167702616681295>.
- Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: ending the imagery debate. *Proceedings of the National Academy of Sciences*, 112(33), 10089–10092. <https://doi.org/10.1073/pnas.1504933112>.
- Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, 46(4), 1023–1031. <https://doi.org/10.3758/s13428-013-0434-y>.
- RStudio Team (2018). RStudio: Integrated Development for R. Boston: RStudio, Inc. <http://www.rstudio.com/>.
- Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 362(1481), 773 LP–773786. <https://doi.org/10.1098/rstb.2007.2087>.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews. Neuroscience*, 8, 657. <https://doi.org/10.1038/nrn2213>.
- Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events. *Annals of the New York Academy of Sciences*, 1124(1), 39–60. <https://doi.org/10.1196/annals.1440.001>.
- Seligman, M. E. P., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science*, 8(2), 119–141. <https://doi.org/10.1177/1745691612474317>.
- Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 1(2), 213–220. <https://doi.org/10.1177/2167702612469015>.
- Standage, H., Ashwin, C., & Fox, E. (2010). Is manipulation of mood a critical component of cognitive bias modification procedures? *Behaviour Research and Therapy*, 48(1), 4–10. <https://doi.org/10.1016/j.brat.2009.08.005>.
- Steinman, S. A., & Teachman, B. A. (2015). Training less threatening interpretations over the internet: does the number of missing letters matter? *Journal of Behavior Therapy and Experimental Psychiatry*, 49, 53–60. <https://doi.org/10.1016/j.jbtep.2014.12.004>.
- Szpunar, K. K., & Schacter, D. L. (2013). Get real: Effects of repeated simulation and emotion on the perceived plausibility of future experiences. *Journal of Experimental Psychology: General*, 142(2), 323–327. <https://doi.org/10.1037/a0028877>.
- Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences*, 111(52), 18414 LP–18418421. <https://doi.org/10.1073/pnas.1417144111>.
- Wickham, R. E., Reed, D. E., & Williamson, R. E. (2015). Establishing the psychometric properties of the Self and Perceived-partner Authenticity in Relationships Scale-Short Form (AIRS-SF): measurement invariance, reliability, and incremental validity. *Personality and Individual Differences*, 77, 62–67. <https://doi.org/10.1016/j.paid.2014.12.049>.
- Wu, J. Q., Szpunar, K. K., Godovich, S. A., Schacter, D. L., & Hofmann, S. G. (2015). Episodic future thinking in generalized anxiety disorder. *Journal of Anxiety Disorders*, 36, 1–8. <https://doi.org/10.1016/j.janxdis.2015.09.005>.
- Wurtele, S. K., Simons, D. A., & Moreno, T. (2013). Sexual interest in children among an online sample of men and women: prevalence and correlates. *Sexual Abuse*, 26(6), 546–568. <https://doi.org/10.1177/1079063213503688>.
- Yiend, J., Lee, J. S., Tekes, S., Atkins, L., Mathews, A., Vrinten, M., et al. (2014). Modifying interpretation in a clinically depressed sample using “cognitive bias modification-errors”: a double blind randomised controlled trial. *Cognitive Therapy and Research*, 38(2), 146–159. <https://doi.org/10.1007/s10608-013-9571-y>.