Episodic specificity induction and scene construction: Evidence for an event construction account

Kevin P. Madore\textsuperscript{a,b,}\(^\ast\), Helen G. Jing\textsuperscript{b}, Daniel L. Schacter\textsuperscript{b}

\textsuperscript{a}Stanford University, Dept. of Psychology, 450 Serra Mall, Stanford, CA 94305, USA
\textsuperscript{b}Harvard University, Dept. of Psychology, 33 Kirkland Street, Cambridge, MA 02138, USA

**Abstract**

Research has suggested that an episodic specificity induction (ESI)– training in recollecting details of a past event– impacts subsequent memory, imagination, problem solving, and creativity. We have hypothesized that induction effects may be attributable to event construction– the assembly and maintenance of a mental scenario filled with setting, people, and action details. We examine whether ESI impacts metrics of event detail in a standard scene construction task, which is a paradigm focused on the spatial integrity of a mental scenario and the stage upon or setting in which such a scenario occurs. Relative to a control, ESI significantly increased details generated across all categories of event detail in scene construction, including spatial references, entities present, sensory descriptions, and thoughts/emotions/actions. ESI did not influence scores on the Spatial Coherence Index, a critical measure of spatial processing. These findings inform theoretical and functional accounts of the nature and malleability of constructive retrieval.

**1. Introduction**

Recent work indicates that episodic memory contributes to a wide range of cognitive functions, such as future event simulation (Atance & O’Neill, 2001; Schacter, 2012; Schacter et al., 2012; Schacter, Benoit, & Szpunar, 2017; Szpunar, 2010; Tulving, 2002), means-end problem solving (Sheldon, McAndrews, & Moscovitch, 2011), and divergent creative thinking (Addis, Pan, Musicaro, & Schacter, 2016; Duff, Kurczek, Rubin, Cohen, & Tranel, 2013). We have attempted to identify the role that episodic memory retrieval may play in such cognitive functions through the use of an episodic specificity induction (ESI): brief training in recollecting details of recent experiences. ESI is based on the Cognitive Interview (Fisher & Geiselman, 1992; Memon, Meissner, & Fraser, 2010), which is a forensic protocol that focuses on detailed episodic retrieval of a specific past event. We predicted that if a cognitive task recruits episodic memory retrieval, then receiving an ESI prior to the task should enhance performance on that task relative to receiving a control induction that does not recruit the retrieval of episodic details. For the ESI in our studies, participants viewed a video of people carrying out activities in a kitchen and were then given the ESI (an adapted version of the Cognitive Interview) that required them to recall in as much detail as possible the contents they had observed in the video. In a control condition, participants viewed a different version of the video and provided their general impressions of how well the video was made, how much they liked it, and so forth. Following the ESI and impressions control induction, participants performed tasks that we hypothesized did or did not rely on episodic retrieval.

Using this basic procedure, it has been found that the ESI selectively impacts performance on subsequent tasks that we hypothesized involve constructive uses of episodic retrieval, including remembering past experiences and imagining future experiences...
occur in past, imagining the future, and related processes involve generating mental events situated in a space, and prompts during the ESI ask events, imagining future events, and related cognitions could be attributable to the targeting of scene construction. Remembering the where and a scene construction account of ESI e characterized by overgeneralized memory such as aging, may impact the e prompts in these sorts of trainings, whether with individuals with psychopathology or individuals from other populations char-

Blackwell, & Dalgleish, 2017

2007; Hassabis, Kumaran, Vann, & Maguire, 2007). According to this view, scene construction is involved in but not restricted to episodic memory and it has been proposed that scene construction could underlie similarities documented in remembering past events, imagining future events, and related cognitions (Mulally & Maguire, 2014; see also, Rubin & Umanath, 2015). In ESI studies, scene construction could potentially be facilitated by retrieval attempts that focus on filling a mental scenario with details in a spatially coherent context that can be drawn from episodic memories or that are atemporal in nature (for further discussion, see Palombo, Hayes, Peterson, Keane, & Verfaellie, 2018; Roberts, Schacter, & Addis, 2018; Sheldon & El-Asmar, 2018).

In the present experiment, we test a further prediction of this event construction account, namely that ESI effects should be observed on particular measures in a task that taps scene construction, a paradigm focused on the spatial coherence of a mental scenario and the stage upon or setting in which such a scenario occurs (Hassabis & Maguire, 2007; Hassabis, Kumaran, & Maguire, 2007; Hassabis, Kumaran, Vann, & Maguire, 2007). According to this view, scene construction is involved in but not restricted to episodic memory and it has been proposed that scene construction could underlie similarities documented in remembering past events, imagining future events, and related cognitions (Mulally & Maguire, 2014; see also, Rubin & Umanath, 2015). In ESI studies, scene construction could potentially be facilitated by retrieval attempts that focus on filling a mental scenario with details in a spatially coherent context that can be drawn from episodic memories or that are atemporal in nature (for further discussion, see Palombo, Hayes, Peterson, Keane, & Verfaellie, 2018; Roberts, Schacter, & Addis, 2018; Sheldon & El-Asmar, 2018).

Recent work from Rubin, Deffler, and Umanath (2019) highlights the role that scene construction may play during tasks for which induction-related effects have been observed. The researchers define a scene as “a place where a real or fictitious event occurs,” and go on “to contrast the spatial organization of [a] scene to its contents (p. 44).” They indicate that scene construction involves “the where” of the contents or details in an event rather than the contents or details themselves. They also state: “Hassabis and Maguire (2007, p. 304) note that memories of events need a ‘stage on which the remembered event is played or the ‘where’ for the ‘what’ to occur in’; that is, they must have spatial organization (p. 46).” Given this framework, prior ESI-related effects on remembering past events, imagining future events, and related cognitions could be attributable to the targeting of scene construction. Remembering the past, imagining the future, and related processes involve generating mental events situated in a space, and prompts during the ESI ask participants to focus on both the ‘where’ and ‘the what’ of an event.

Adjudicating between a scene construction vs. event construction account of facilitative effects of ESI is motivated both theoretically and functionally. Theoretically, understanding the mechanism underlying ESI effects may be useful for identifying the boundary conditions and nature of episodic memory more broadly (Addis, 2016; Conway & Lovejoy, 2015; de Vito, Gambo, & Brandimonte, 2012; Eichenbaum, 2017; Moscovitch, Cabeza, Wincour, & Nadel, 2016; Schiller et al., 2015). Functionally, understanding the mechanism underlying ESI effects may be useful for building on therapeutic interventions like MEST (Memory specificity Training; Raes, Williams, & Hermans, 2009) that have been adopted to boost specificity and reduce symptoms of psychopathology in depression and posttraumatic stress disorder (for review, see Erten & Brown, 2018; Hitchcock, Werner-Seidler, Blackwell, & Dalgleish, 2017; and Halldorf, Austin, Takano, & Raes, 2018). Placing an emphasis on event vs. scene construction prompts in these sorts of trainings, whether with individuals with psychopathology or individuals from other populations characterized by overgeneralized memory such as aging, may impact the efficacy of training.

We think that the standard scene construction task can be used to adjudicate between an event construction account of ESI effects and a scene construction account of ESI effects because particular metrics within the task assay strictly spatial vs. event processing. By spatial, we mean ‘the where’ of a construction, and by event we mean the ‘what’ of a construction. In the scene construction procedure (Hassabis, Kumaran, Vann, et al., 2007), participants verbalize constructed scenes from atemporal cues (e.g., “Imagine you’re lying on a deserted white sandy beach...”) and then rate their constructions for various phenomenological characteristics, such as vividness/salience, sense of presence, and spatial coherence. Scorers then code participants’ narratives for details from four categories of content, including spatial references, entities present, sensory descriptions, and thoughts/emotions/actions, as well as provide an overall quality judgment of the vividness and richness of the imagined scene. The prediction that the ESI affects the construction of mental events on downstream tasks suggests that the key ingredient of ESI may be the necessity of generating detailed mental events. Such effects should be evident on several measures of event detail that are assessed in the standard scene construction task, including details from the specific categories of spatial references, entities present, sensory descriptions, and thoughts/emotions/actions.
The concepts of event construction and scene construction are not identical because scene construction places special emphasis on the spatial qualities of a mental scenario and the stage upon or setting in which a scenario transpires, that is, spatial organization. These spatial qualities are more directly assessed by the Spatial Coherence Index, a key measure of the scene construction task focused on “the contiguousness and spatial integrity of the imagined scene” (Hassabis, Kumaran, Vann, et al., 2007, p. 1728). In an earlier characterization of the effects of the specificity induction, Schacter and Madore (2016, pp. 250–251) noted substantial overlap between the concepts of “scene construction” and “event construction,” but nonetheless drew a distinction between them, noting that: “one of the key measures of scene construction, the spatial coherence index, assesses the spatial integrity of a constructed scene (Hassabis, Kumaran, Maguire, 2007; Hassabis, Kumaran, Vann, et al., 2007). We do not claim that the specificity induction selectively impacts the spatial coherence of a constructed scene. Instead, we suggest that the induction could potentially impact the details associated with both elements of a scene and their relations, including spatial relations.” That is, we predicted that ESI effects in the standard scene construction task would be exhibited in ‘the what’ of event details rather than in ‘the where’ of spatial organization or coherence.

Here we attempt to distinguish between the two concepts by taking a narrow view of spatial processing in the scene construction task: only the scene construction account predicts an effect on the Spatial Coherence Index in particular because this metric is narrowly focused on spatial processing. To adjudicate between an event construction account of ESI effects and a scene construction account of ESI effects, we provide two plausible accounts of potential data patterns. If ESI affects primarily an event construction process, then we should observe effects of ESI on detail content measures of scene construction (i.e., ‘the what’). On the contrary, if ESI affects primarily a scene construction process, then we should observe effects of ESI on the Spatial Coherence Index in particular (i.e., ‘the where’).

To address these issues, in the current study we adopted a one-session experimental design in which participants received either the specificity induction or a control induction not focused on specific retrieval before completing scene construction prompts and then switched to receiving the second induction before completing additional scene construction prompts. We adopted the procedure and scoring of Hassabis, Kumaran, Vann, et al. (2007), and analyzed all unweighted scene construction metrics, including (a) vividness/salience, (b) sense of presence, and (c) spatial coherence from participants, and (d) detail content and (e) quality judgment ratings from scorers.

The standard scene construction task also includes an overall Experiential Index score that rescales and weights these five metrics to obtain a summated value of imagined experience on the task with four subcomponents (vividness/salience and sense of presence; spatial coherence, detail content, and quality judgment). According to Hassabis, Kumaran, Vann, et al. (2007), each subcomponent represents a different facet of performance on the scene construction task. The vividness/salience and sense of presence subcomponent measures the key phenomenological quality of each construction; the spatial coherence subcomponent measures the extent to which each construction evokes an integrated spatial context; the detail content subcomponent measures the most directly observable information about each construction with separated and classified segments; and the quality judgment subcomponent measures the overall ‘picture’ of each construction. We thus also rescaled all scene constructions for this weighted score from the subcomponents. This weighted score represents both spatial and event processing in the scene construction task (see Material and Methods), or the overall mental experience rather than spatial scene. We were motivated to analyze the overall Experiential Index for two reasons. First, the researchers who devised the scene construction task (Hassabis, Kumaran, Vann, et al., 2007) note that the index weights the most directly observable information the most (i.e., detail counts from content categories). For methodological rigor, we thought it was important to examine participants’ scene constructions with these weightings. Second, we included the Experiential Index because it represents the overall experience of scene construction rather than just the spatial organization aspect of scene construction; we reasoned that ESI would affect this index of overall experience not restricted to spatial coherence if an event construction account is viable.

To anticipate the results, participants generated significantly more details across all categories of scored content following the specificity induction relative to the control, but did not exhibit significant differences in scores on the Spatial Coherence Index as a function of induction. These results were observed whether unweighted or weighted metrics of scene construction were analyzed. The weighted Experiential Index score was also significantly increased following the specificity induction relative to the control, and this increase was driven by the weighted details from the content categories, as well as the weighed quality judgment.

2. Material and methods

2.1. Participants

Twenty-four young adults ($M_{\text{age}} = 21.42$ years, $SD_{\text{age}} = 2.86$, range$_{\text{age}} = 18–27$, 16 female) completed the study and were recruited from advertisements at Boston University and Harvard University. They provided written consent and received pay for participation, and were treated in a manner approved by Harvard University’s ethics committee. All participants had normal or corrected-to-normal vision and no current or former neurological impairment. One additional participant was excluded for task noncompliance. We selected a sample size of 24 useable subjects and stopped data collection after reaching this number because an a priori power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that 24 was an appropriate number for detecting a medium-sized effect ($d = 0.60$, power > 0.80, two-tailed for a within-subjects design). Our prior ESI-related studies (e.g., Madore et al., 2014, 2015) have also adopted this sample size and detected medium- to large-sized effects.
Segment 1: (~45mins)
Induction phase:
2-min video + 2-min filler + questions about the video (episodic specificity induction or impressions control)

Task phase:
Untimed scene construction prompts and ratings

Segment 2: (~45mins)
Induction phase:
2-min video + 2-min filler + questions about the video (episodic specificity induction or impressions control)

Task phase:
Untimed scene construction prompts and ratings

Fig. 1. Schema of experimental design. Note. Order of inductions was counterbalanced between participants. Order of scene construction prompts (from two sets of 5 cues) was also counterbalanced between inductions. A 5-minute math filler task came between segments 1 and 2. Task scoring included all unweighted and weighted scene construction metrics.

2.2. Materials, design, and procedure

2.2.1. Overview

As depicted in Fig. 1, participants came to the lab for one session and completed two main segments. In each segment, which lasted approximately 45 min, participants first watched a 2-minute video of a man and woman performing activities in a kitchen (time sliced from materials in Koutstaal, Schacter, Johnson, Angell, & Gross, 1998), completed a 2-minute math filler task (addition and subtraction), and then received the episodic specificity induction or an impressions control induction with questions about the content of the video they had viewed. After each induction, participants completed five scene construction trials. A 5-minute math filler task (addition and subtraction) was used between the end of the first segment and the beginning of the second segment (as in other studies with single-session induction paradigms; see Madore & Schacter, 2016, for an example). Different stimuli were used in each segment (e.g., video, induction, and scene construction prompts). One experimenter ran the induction portions of the experiment with each participant, and a separate experimenter blind to induction received the scene construction portions with each participant to minimize potential biases. Participants verbalized their answers to the induction and the scene construction prompts. None of the 24 participants reported awareness of the experimental manipulation or hypotheses in debriefing.

There are a few aspects of our general methodological design that should be noted, namely our selection of a single-session design with two segments, and our control induction. We have found indistinguishable ESI-related behavioral and neural effects in all of our previous studies whether a single-session or multi-session protocol is adopted (i.e., two inductions in one session or one induction in two separately spaced sessions), as well as whether the impressions control or a math control is compared to ESI (for examples, see Madore et al., 2014; Madore, Szpunar, Addis, & Schacter, 2016). We do not think that these aspects of the methodology are important or that they impact our results given existing evidence on ESI and related manipulations (e.g., construal level theory, Trope & Liberman, 2010) where brief within-session and across-session manipulations lead to similar effects.

2.2.2. Inductions

Participants were randomly assigned to a video-induction sequence (which was counterbalanced across participants). For the episodic specificity induction, based on the Cognitive Interview (Fisher & Geiselman, 1992; Memon et al., 2010), participants completed three mental imagery probes about the setting, people, and actions of the video they had watched. They were asked to close their eyes and get a picture in their head about the respective aspect of the video and were instructed to report everything and be as detailed as possible, as they were the chief expert. Open-ended follow-up probes were used to elicit additional details about aspects of the video the participants had discussed. For the impressions control induction, participants were also asked to reflect back on and report about aspects of the video they had seen but not in an episodically specific way. Participants were first asked about their general impressions and thoughts of the video, and were then guided to generate additional information about the video using probes from a question bank. For example, participants were asked to generate adjectives that described the setting, people, and actions, and to estimate when and how the video was made. No mental imagery probes were used. After answering the questions from the bank, participants were then asked if they had any additional or concluding thoughts about the video. Interviews were approximately 5 min each, and did not differ significantly in length as a function of induction (for full scripts, see Madore et al., 2014).

2.2.3. Scene construction task

After completing the induction phase in each segment, participants completed the scene construction task with the instructions, materials, and procedures used by Hassabis, Kumaran, Vann, et al. (2007). Participants were presented with 5 individual prompts in each segment that cued a different scene. They were instructed to imagine a vivid scene in their mind’s eye related to each respective prompt, to give free rein to their imagination and not recount a memory, and to be as detailed as possible, including information related to all available senses (e.g., sight, sound, and smell). Participants were provided with an example prompt and response to
ensure comprehension. The experimenter sat facing each participant, read each of the 5 prompts aloud one at a time, and gave each participant a note card on which the respective prompt was delineated (e.g., “Imagine you’re lying on a deserted white sandy beach in a beautiful tropical bay. I want you to describe the experience and the surroundings in as much detail as possible using all your senses including what you can see, hear, and feel.”). Participants had unlimited time to verbalize their response and general probing was used by the experimenter (e.g., “Is there anything else you can tell me?”) until participants came to a natural end. Note that because the scene construction task includes an example prompt and response, and participants are cued to report as much detail as possible before each experimental trial, any effects of the specificity induction observed are above and beyond report criterion, decision threshold, or response biasing effects. After each trial, participants completed 5 ratings focused on different aspects of their construction (from 1 = least to 5 = most) on a sheet of paper provided to them: difficulty, vividness/salience, detail, sense of presence, and similarity to past memories. Upon completing these ratings, participants filled out the Spatial Coherence Index on a separate sheet of paper provided to them. They viewed the 12 items of the index and checked off which one(s) best described their imagined scene and experience of imagining it. After filling out the Spatial Coherence Index, participants were asked if they had described a memory or created a new story. Participants then completed the next scene construction prompt. It should be noted that Hassabis, Kumaran, Vann, et al. (2007) used 10 scene construction prompts per participant; in the current study we used 5 different scene construction prompts per participant in each induction segment to compare effects of the manipulation on performance. The prompts were blocked into two sets of 5 (Block 1: beach, museum, pub, weekend, and Christmas and Block 2: ship, street market, forest, castle, and friend) and counterbalanced with induction order across participants. While some prompts were atemporal in nature and some involved episodic future thinking, we collapsed performance across prompt type for the main analyses (as in Hassabis, Kumaran, Vann, et al., 2007) because there were no demonstrable differences as a function of this variable.

2.3. Scoring and coding

The task trials were audio-recorded and later transcribed. We adopted the scoring and coding measures of Hassabis, Kumaran, Vann, et al. (2007) to test for effects of induction on particular aspects of scene construction performance. Below we outline the different metrics of scene construction, both unweighted values and then weighted values used in the Experiential Index, and highlight those that are most narrowly implicated in spatial or scene processing.

As is typical in scene construction work (e.g., Hassabis, Kumaran, Vann, et al., 2007), four cognitive subcomponents were obtained to measure task performance. Two subcomponents came from ratings by participants during the tasks. The additional two subcomponents came from ratings by one of two coders blind to induction and to all experimental hypotheses scored after the tasks were completed.

From participants, one subcomponent of performance was ratings given for vividness/salience (from 1 to 5) and sense of presence per trial (from 1 to 5) for the global scene. Another subcomponent was ratings given on the Spatial Coherence Index per trial. Of the 12 items, 8 contributed to a spatially integrated response (e.g., “I could see it as one whole scene in my mind’s eye.”) and 4 to a spatially fragmented response (e.g., “It wasn’t so much a scene as a collection of images.”). One point was added for each integrated response and one point was subtracted for each fragmented response. Participants could receive a possible score of −4 to 8 per trial. Of these metrics, the Spatial Coherence Index is the narrowest measure of scene processing because it specifically refers to space, whereas the vividness/salience and sense of presence ratings refer to the global scene (which could include non-spatial or event aspects). These metrics represent unweighted values before rescaling for the Experiential Index.

From blind coders, one subcomponent of performance was a quality judgment of each overall scene construction provided by participants per trial. Coders were asked to get a picture in their mind’s eye from each response and rate it from 0 to 10 in terms of detail quality. Before scoring the experimental trials, two blind coders received training independently and scored an interrater packet of 10 responses from a pilot dataset with high reliability (Cronbach’s alpha = 0.91). Each rater scored half of the experimental trials.

Another subcomponent of performance was content, a sum of details contained in each scene construction per trial from the specific categories of spatial references, entities present, sensory descriptions, and thoughts/emotions/actions (see Hassabis, Kumaran, Vann, et al., 2007, for examples of details from each category). The two blind coders segmented each construction into different bits of information and then labeled each detail as fitting into one of the categories as appropriate. Details that could not fit into one of these categories were not scored (e.g., repetitive statements, off-topic information). Before scoring the experimental trials, the two blind coders received training independently and scored an interrater packet of 10 responses from a pilot dataset with high reliability (Cronbach’s alpha = 0.94 for spatial references, 0.93 for entities present, 0.90 for sensory descriptions, and 0.93 for thoughts/emotions/actions). Each rater scored half of the main trials. These metrics also represent unweighted values before rescaling for the Experiential Index.

The four unweighted subcomponents of performance were obtained to assess aspects of scene construction performance. As in Hassabis, Kumaran, Vann, et al. (2007), the four subcomponents were then weighted and summed to create an overall Experiential Index (ranging from 0 to 60 points or not experienced at all to richly experienced). Rescaling was done to obtain the overall Experiential Index score (see Hassabis, Kumaran, Vann, et al., 2007). Ratings of vividness/salience and sense of presence were each rescaled from 1 to 5 to 0 to 4 points (6.5% of the overall score per rating). Ratings on the Spatial Coherence Index were normalized and rescaled around 0 to give a score between −6 and 6, and participants were only included if above 0 points (10% of the overall score). Of the 24 participants, 22 had Spatial Coherence Index scores above 0 points; induction-related effects were not dependent on whether data from the 22 participants or all 24 participants were analyzed. Quality judgment scores were rescaled from 0 to 10 to 0 to 18 points (30% of the overall score). Detail content included a maximum of 7 details per four categories per trial for a total of 0 to 28
points (47% of the overall score). These metrics represent the weighted values for the Experiential Index.

We report below the results for all unweighted and weighted metrics, and also highlight those narrowly implicated in spatial processing, including the Spatial Coherence Index. All data are available upon request.

3. Results

In terms of analytic checks, all trials for each of the 24 participants were included for the unweighted metrics. As noted above, all trials for the 22 participants with Spatial Index Coherence scores above 0 were included for the weighted metrics. There were no significant effects of the between-subjects variable of induction order/carryover effects on results for any metrics (ps ≥ .20).1 Word count per trial also did not significantly vary as a function of induction, $F(1, 23) = 1.97, \text{MSE} = 1879.62, p = .174, \eta^2_p = .08$. Participants generated 203.25 words per scene construction trial ($\text{SE} = 23.50$) following the control induction and 220.80 words per trial ($\text{SE} = 22.11$) following the specificity induction. This finding indicates that overall verbosity can be ruled out as an explanation for any induction-related effects.

To assess impacts of the specificity induction on particular metrics of scene construction, we conducted a series of repeated-measures ANOVAs with the within-subjects factors of induction (control vs. specificity) and the respective dependent variable(s) of interest, starting with each unweighted subcomponent of scene construction. We highlight the Spatial Coherence Index (as done by Hassabis, Kumaran, Vann, et al. (2007)). We then repeat the same analyses but incorporate each weighted subcomponent of scene construction for the Experiential Index, and again highlight the Spatial Coherence Index. We report the results of the models below, moving from main effects to two-way interactions where appropriate (Greenhouse-Geisser corrections were used for sphericity violations).

3.1. Unweighted scene construction metrics

The descriptive statistics of the unweighted results are presented in Table 1 below. We first examined induction-related effects on each unweighted subcomponent of scene construction: vividness/salience ratings, sense of presence ratings, and Spatial Coherence Index by participants, and quality judgments and detail content by blind raters.

We first examined induction-related effects on each unweighted subcomponent of scene construction: vividness/salience ratings, sense of presence ratings, and Spatial Coherence Index by participants, and quality judgments and detail content by blind raters.

Vividness/salience. There was no significant main effect of induction on unweighted vividness/salience ratings, $F(1, 23) < 1$, MSE = 0.10, $p = .66, \eta^2_p = .01$.

Sense of presence. There was also no significant main effect of induction on unweighted presence ratings, $F(1, 23) < 1$, MSE = 0.13, $p = .58, \eta^2_p < .01$.

Spatial Coherence Index. There was a marginal effect of induction on the unweighted Spatial Coherence Index in the opposite direction of that predicted, $F(1, 23) = 3.62, \text{MSE} = 0.72, p = .07, \eta^2_p = .14$. Participants actually exhibited a marginally higher score on the Spatial Coherence Index following the control induction compared with the specificity induction.2

Quality judgment. There was a significant main effect of induction on the unweighted quality judgment rating, $F(1, 23) = 13.96, \text{MSE} = 0.23, p = .001, \eta^2_p = .38$. Quality of scene construction was rated as higher following the specificity induction compared with the control.

Detail content. There was also a significant main effect of induction on the unweighted number of details generated across categories of content, $F(1, 23) = 18.72, \text{MSE} = 11.67, p < .001, \eta^2_p = .45$. A significantly greater number of details were generated following the specificity induction relative to the control induction within each category of content, including spatial references, entities present, sensory descriptions, and thoughts/emotions/actions. There was also a significant main effect of category on number of details, $F(1.97, 45.38) = 54.88, \text{MSE} = 27.32, p < .001, \eta^2_p = .71$, such that more details were generated for entities present relative to sensory descriptions relative to thoughts/emotions/actions relative to spatial references (ps ≤ .02). Critically, the interaction of induction and detail content was non-significant, $F(1.91, 43.96) = 0.58, \text{MSE} = 7.33, p = .63, \eta^2_p = .03$. This result indicates that the specificity induction increased the number of details that participants generated from all categories of information to a similar degree.

Taken together, the results from the unweighted subcomponent analyses indicate that the specificity induction increased quality judgments and the number of details within each category of content of scene construction and did not increase ratings of vividness, sense of presence, or spatial coherence. These results indicate that specificity induction effects are not restricted to spatial processes because the manipulation affected both spatial and non-spatial elements of a mental scene and did not affect spatial coherence. While

---

1 For completeness, we report p-values for induction x carryover effects on all dependent variables in the same order as presented in Results for unweighted metrics: vividness/salience rating ($p = .54$), sense of presence rating ($p = .81$), Spatial Coherence Index ($p = .60$), spatial reference details ($p = .31$), entities present details ($p = .20$), sensory description details ($p = .92$), thought/emotion/action details ($p = .73$), quality judgment rating ($p = .56$), difficulty rating ($p = .46$), detail rating ($p = .73$), and similarity rating ($p = .21$). In addition, there were no induction x carryover effects observed for weighted metrics (ps ≥ .48), including the Experiential Index ($p = .57$).

2 For completeness, we also analyzed Spatial Coherence Index for the 12 participants who received the control induction first vs. the 12 participants who received the specificity induction first. There was no significant effect of induction on Spatial Coherence Index with this between-subjects analysis ($p = .17$ favoring control induction).
After examining induction-related effects on the weighted subcomponents of scene construction, we assessed the effect of the specificity manipulation on the weighted subcomponents of scene construction using the rescaling criteria from Hassabis, Kumaran, Vann, et al. (2007) for vividness/salience, sense of presence, and Spatial Coherence Index from participants and quality judgment and detail content from blind scorers. To anticipate the results, the same induction-related effects were observed for weighted vs. unweighted subcomponents. We also report induction-related scores on the Experiential Index using the weighted subcomponents. As noted above, 2 participants did not receive Spatial Coherence Index scores above 0 so their data are not reported below (see Hassabis, Kumaran, Vann, et al., 2007). Induction-related effects for weighted subcomponents did not vary based on whether these participants were included or excluded from analyses.

Vividness/salience. There was no significant effect of induction on weighted vividness/salience ratings, $F(1, 21) < 1$, $MSE = 0.10$, $p = .707$, $\eta^2_p = 0.01$.

Sense of presence. There was also no significant effect of induction on weighted sense of presence ratings, $F(1, 21) < 1$, $MSE = 0.13$, $p = .864$, $\eta^2_p < 0.01$.

Spatial Coherence Index. There was a marginal effect of induction on the weighted Spatial Coherence Index favoring the control induction rather than the specificity induction, $F(1, 21) = 2.82$, $MSE = 0.78$, $p = .018$, $\eta^2_p = 0.12$.

Quality judgment. There was a significant effect of induction on the weighted quality judgment favoring the specificity induction relative to the control, $F(1, 21) = 14.69$, $MSE = 0.75$, $p = .001$, $\eta^2_p = 0.41$.

Detail content. As with the unweighted results, there were significant main effects of induction, $F(1, 21) = 29.88$, $MSE = 0.87$, $p < .001$, $\eta^2_p = 0.59$, and quality of weighted content, $F(3, 63) = 39.33$, $MSE = 1.74$, $p < .001$, $\eta^2_p = 0.65$. For the main effect of induction, more details were generated across all categories of content following the specificity induction relative to control. For the main effect of category of weighted content, more details were generated for entities present relative to sensory descriptions relative to spatial references and thoughts/emotions/actions ($ps \leq 0.008$); details for spatial references and thoughts/emotions/actions did not significantly vary ($p = .12$). Critically, the interaction of induction and category of detail was non-significant, $F(2, 01, 42.12) = 2.23$, $MSE = 1.09$, $p = .119$, $\eta^2_p = 0.10$. This finding indicates that more details were generated for all categories of weighted content following the specificity induction relative to control. For thoroughness (see Table 2), we note that post-hoc tests indicated that induction-related effects were driven by the categories of spatial references, $F(1, 21) = 14.61$, $MSE = 0.82$, $p = .001$, $\eta^2_p = 0.41$; sensory descriptions, $F(1, 21) = 13.55$, $MSE = 0.80$, $p = .001$, $\eta^2_p = 0.39$; and thoughts/emotions/actions, $F(1, 21) = 5.99$, $MSE = 1.26$, $p = .023$, $\eta^2_p = 0.22$. The induction-related effect for entities present was not significant, $F(1, 21) = 2.82$, $MSE = 0.17$, $p = .108$, $\eta^2_p = 0.12$.

Experiential Index (EI). There was a significant effect of induction on this overall score, $F(1, 21) = 14.23$, $MSE = 9.84$, $p = .001$, $\eta^2_p = 0.40$. Participants exhibited a higher overall Experiential Index score following the specificity induction compared with the control, which was driven by the quality judgment and detail content subcomponents of performance. These results point to effects of the specificity induction on the overall mental experience, or aspects of performance that are not limited to spatial coherence or spatial processing of a mental scene.

### 3.2. Weighted scene construction metrics

<table>
<thead>
<tr>
<th>Subcomponents</th>
<th>Control induction</th>
<th>Episodic specificity induction</th>
<th>Induction-related mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vividness/salience (1–5) by participant</td>
<td>3.30 (0.10)</td>
<td>3.26 (0.11)</td>
<td>−0.04 (0.09)</td>
</tr>
<tr>
<td>Sense of presence (1–5) by participant</td>
<td>3.21 (0.09)</td>
<td>3.15 (0.12)</td>
<td>−0.06 (0.10)</td>
</tr>
<tr>
<td>Spatial coherence index (−4 to 8) by participant</td>
<td>4.20 (0.35)</td>
<td>3.73 (0.32)</td>
<td>−0.47 (0.25)</td>
</tr>
<tr>
<td>Quality judgment (0–10) by blind rater</td>
<td>5.99 (0.23)</td>
<td>6.50 (0.24)</td>
<td>0.52 (0.14)</td>
</tr>
</tbody>
</table>

**Number of details:**
- Spatial references: 3.61 (0.43) vs. 5.13 (0.48) $p = .04$.
- Entities present: 13.63 (1.17) vs. 16.06 (1.34) $p = .002$.
- Sensory descriptions: 7.45 (1.23) vs. 10.03 (1.59) $p = .01$.
- Thoughts/emotions/actions by blind rater: 5.43 (0.63) vs. 7.43 (0.97) $p = .03$.

**Note.** Numeric values are presented as mean per trial (with standard error in parentheses). The induction-related mean difference per trial (with standard error of the mean difference in parentheses) is presented in the last column.

* Indicates that $p < .05$.

** Indicates that $p < .01$.

As with the unweighted Spatial Coherence Index, the same pattern of results was obtained when analyzing the 12 participants who received the control induction first vs. the 10 participants who received the specificity induction first, ($p = .54$, favoring the control induction).
### Table 2
Descriptive statistics for weighted subcomponents on scene construction task as a function of induction, followed by overall score.

<table>
<thead>
<tr>
<th>Weighted scene construction metrics (least to most)</th>
<th>Control induction</th>
<th>Episodic specificity induction</th>
<th>Induction-related mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subcomponents</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vividness/salience (0–4) by participant, 6.5% of EI</td>
<td>2.34 (0.10)</td>
<td>2.31 (0.11)</td>
<td>–0.04 (0.10)</td>
</tr>
<tr>
<td>Sense of presence (0–4) by participant, 6.5% of EI</td>
<td>2.21 (0.09)</td>
<td>2.18 (0.13)</td>
<td>–0.03 (0.11)</td>
</tr>
<tr>
<td>Spatial coherence index (0–6) by participant, 10% of EI</td>
<td>2.41 (0.34)</td>
<td>1.96 (0.30)</td>
<td>0.45 (0.27)</td>
</tr>
<tr>
<td>Quality judgment (0–18) by blind rater, 30% of EI</td>
<td>10.85 (0.45)</td>
<td>11.85 (0.43)</td>
<td>1.00 (0.26)</td>
</tr>
<tr>
<td><strong>Number of details (capped at 7)</strong>:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial references</td>
<td>3.41 (0.30)</td>
<td>4.45 (0.27)</td>
<td>1.05 (0.27)</td>
</tr>
<tr>
<td>Entities present</td>
<td>6.68 (0.10)</td>
<td>6.89 (0.07)</td>
<td>0.21 (0.12)</td>
</tr>
<tr>
<td>Sensory descriptions</td>
<td>4.85 (0.38)</td>
<td>5.85 (0.27)</td>
<td>0.99 (0.27)</td>
</tr>
<tr>
<td>Thoughts/emotions/actions</td>
<td>4.05 (0.34)</td>
<td>4.87 (0.29)</td>
<td>0.83 (0.33)</td>
</tr>
<tr>
<td>Summed detail (capped at 28)</td>
<td>18.99 (0.79)</td>
<td>22.06 (0.65)</td>
<td>3.07 (0.56)</td>
</tr>
<tr>
<td>by blind rater, 47% of EI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall Experiential index (0–60)</strong></td>
<td>36.80 (1.31)</td>
<td>40.37 (1.23)</td>
<td>3.56 (0.94)</td>
</tr>
</tbody>
</table>

Note. Numeric values are presented as mean per trial (with standard error in parentheses). The induction-related mean difference per trial (with standard error of the mean difference in parentheses) is presented in the last column. EI = Experiential Index.

* Indicates that p < .05.

** p < .01.

*** p ≤ 0.001.

### 4. Discussion

Here we tested whether an ESI impacts particular metrics of scene construction, a task focused on spatial aspects of a mental scenario. Consistent with predictions from an event construction perspective, participants provided more details across all categories of unweighted and weighted content following the ESI relative to the control (i.e., ‘the what’). However, unweighted and weighted scores on the Spatial Coherence Index – considered to be a key measure of scene construction – did not show significant induction-related differences (i.e., ‘the where’); in contrast, there were actually some trending induction effects in the opposite pattern to that predicted. We thus extend the effects of ESI to metrics within a novel task domain that draws on event construction processes. These results are in line with our previous induction-related studies (see Schacter & Madore, 2016, for review) where effects of the manipulation have been observed on tasks that include but are not limited to spatial processing, such as remembering past and imagining future events, solving problems, and thinking creatively. Most relevant to the current study, prior work (e.g., Madore et al., 2014) has found that induction-related effects are typically exhibited in terms of total internal or episodic details produced on generative tasks, in line with the observed effects in terms of total event details across categories of content whether unweighted or weighted.

Why does ESI affect detail generation measures of scene construction? Our results provide evidence in support of Schacter and Madore (2016) suggestion that ESI impacts an event construction process: ESI affected details generated for mental scenes including those from the category of spatial references but also from the categories of entities present, sensory descriptions, and thoughts/emotions/actions. These findings complement the results of Madore et al. (in press) showing that memory specificity and imagination specificity inductions have comparable downstream effects on subsequent memory and imagination tasks, indicating constructive rather than reproductive retrieval is the mechanism underlying ESI effects. Note that our previous research has shown that the ESI does not increase the production of all details on subsequent tasks. For example, although participants provide more internal or episodic details on constructive memory and imagination tasks following ESI vs. a control induction, they do not provide more external or semantic details (e.g., Madore et al., 2014; Madore & Schacter, 2016). And as noted earlier, a prior ESI has no effect on the production of details on a picture description task that does not involve constructive retrieval (in press; Madore et al., 2014). Thus, the present results extend the domain of tasks on which ESI produces an increase in event details during constructive retrieval.

There are a few methodological points worth considering in regard to the pattern of findings. The first involves potential demand characteristics or response biasing as an explanation for ESI-related effects. We do not think that these factors can explain our results because none of the 24 participants in debriefing were aware of the experimental manipulation or hypotheses. Moreover, two experimenters were used in the study. The first ran the induction with the participant and the second – who was blind to which induction had been received – ran the scene construction task, so it seems unlikely that participants would exhibit demand characteristics between two different experimenters. Moreover, receiving instructions to be as detailed as possible cannot account for our results because participants were cued with this instruction before every scene construction trial. Along these lines, prior work has also demonstrated that simply adopting a specific response format cannot account for ESI-related findings on generative tasks: as noted earlier, behavioral ESI effects are not observed when describing pictures in specific detail (in press; Madore et al., 2014), and mixed-effects models show that induction and picture description independently contribute to generation on memory and imagination tasks (Madore et al., 2014). Neuroimaging studies have also found that ESI effects are not observed in the brain or in behavior when generating detailed definitions of objects (Madore, Szpunar, et al., 2016) or associates of them (Madore, Thakral, Beaty, Addis, & Schacter, 2017). The response format (whether verbally, written, or typed), and the length of time spent on each trial (whether self-
paced, three minutes, or five minutes), also cannot explain ESI-related findings; similar effects have been observed irrespective of these factors (e.g., Madore et al., 2014, 2015).

Another methodological point that we can address is the use of a video-induction procedure with a salient setting. The Spatial Coherence Index in particular asks participants to judge the vividness and detail of the scene they have just created in their mind’s eye with 12 different items. It is possible that participants who receive the ESI, which is focused on a specific setting as well as specific people and actions, may be primed to think that they are less vivid in their scene constructions than in their recently generated ESI-dependent memory. This sort of priming could skew Spatial Coherence Index scores in comparison to an impressions control induction where emphasis is not placed on specificity of setting. We think that this interpretation can be ruled out because participants did not exhibit differences on global vividness/salience or detail ratings of scenes as a function of induction, whether specificity or control. While our study was well powered to observe within-subject, medium-sized effects, a future study with a larger sample size may be warranted to investigate potentially small induction-related differences on particular metrics, including the Spatial Coherence Index.

A third methodological consideration that emerged from these studies is that the induction appeared to primarily affect metrics from blind scorers (e.g., detail generation) rather than metrics from participants (e.g., Spatial Coherence Index). In the context of scene construction, it has recently been suggested (Miloyan & McFarlane, 2018) that measurement error can occur when relying on metrics from scorers vs. metrics from participants, in that metrics from scorers may not reflect the actual richness of scene constructions as experienced by participants. We think that brain imaging has shown how metrics from scorers, like detail generation, have utility in capturing the richness of scene construction performance. For example, previous functional magnetic resonance imaging (fMRI) studies (e.g., Madore, Szpunar, et al., 2016) have found that the number of episodic details contained in verbalized constructions as scored by blind raters can modulate the strength of neural activity in key brain regions implicated in these sorts of constructions by participants, such as the hippocampus, inferior parietal lobule, and precuneus (Benoit & Schacter, 2015). Critically, induction-related behavioral differences in episodic details have been found to modulate induction-related neural differences in these brain regions with fMRI (Madore, Szpunar, et al., 2016). Transcranial magnetic stimulation (TMS) has also been utilized in the context of scene and event constructions; inhibiting activity in key regions implicated in generating events, such as the angular gyrus, has been found to decrease verbal detail generation— in particular episodic detail generation— during remembering and imagining events (Bonnici, Cheke, Green, FitzGerald, & Simons, 2018; Thakral, Madore, & Schacter, 2017). Verbal descriptions scored by blind raters, then, appear to be a valid metric of performance on generative construction tasks that are malleable to various experimental manipulations.

A related point raised by Miloyan and McFarlane (2018) is that metrics from participants, such as the Spatial Coherence Index, may rely on retrospective subjective assessment and metacognition over the previously verbalized scene construction. We agree that this approach has limitations, and further note that phenomenological ratings by participants often do not capture the richness of these sorts of constructions as captured by blind scorers or by neuroimaging evidence. For example, older adults often rate their constructions as no less vivid or even more vivid than those of young adults, yet scorers routinely code older adults’ constructions as containing significantly less episodic detail than young adults’ constructions (e.g., Addis, Musicaro, Pan, & Schacter, 2010; Cole, Morrison, & Conway, 2013). Brain imaging has also provided evidence bearing on the limitations of subjective ratings, as older adults typically rate their constructions as no less vivid than those of young adults, yet show less activation in neural regions implicated in detailed event and scene building than young adults do (e.g., Addis, Roberts, & Schacter, 2011). In following the guidelines of Miloyan and McFarlane (2018) for measurement error in these sorts of generative paradigms, we acknowledge that there may be differences in the types of processes assessed by metrics from scorers vs. metrics from participants. We also note that scorers were blind to induction/group membership as well as hypotheses in the current study and attained high inter-rater reliability across all values measured, which should minimize effects of measurement error due to scorer characteristics. While speculative, induction-related differences on metrics from scorers but not metrics from participants also fits with emerging behavioral and neural data (e.g., Kyung, Yanes-Lukin, & Roberts, 2016; Qin, van Marle, Hermans, & Fernandez, 2011; Richter, Cooper, Bays, & Simons, 2016; Sheldon & El-Asmar, 2018) illustrating the ways in which different facets of episodic retrieval and construction can be systematically teased apart, and thus should continue to be explored further in future work.

5. Conclusions

Overall, our results indicate that ESI targets an event construction process that can be identified on standard scene construction metrics. These results are compelling and significant for both theoretical and functional reasons. Conceptually, effects of ESI on subsequent event detail within scene construction fits with emerging frameworks and data that highlight the roles of both spatial and non-spatial or event processing in generative “memory-like” tasks (e.g., Addis, 2018; Conway & Loveday, 2015; de Vito et al., 2012; Eichenbaum, 2017; Moscovitch et al., 2016; Schiller et al., 2015). Functionally, the results may also be important for understanding better the role that event construction plays in mnemonic disruptions in specific populations, such as depression and aging, characterized by overgeneralized memory and related processes (for review, see Erten & Brown, 2018; Hitchcock et al., 2017; and Hallford et al., 2018; for example empirical work, see King et al., 2011). Incorporating event construction prompts in specificity trainings like MEST may boost specificity and reduce symptoms of psychopathology further than current approaches, which could impact the efficacy of training. We suggest that event construction is a malleable and adaptive process (Schacter, 2012) involved in tasks that are not typically viewed as “episodic remembering” but that may rely on the constructive retrieval of details and the assembly and maintenance of a mental event for completion.
Acknowledgements

This research was supported by National Institute of Mental Health grant MH060941 to DLS. KPM is supported by National Institute on Aging grant F32AG059341. We thank Eleanor Maguire for providing the scene construction materials used. We also thank Jyotika Bindra, Haley Dodds, Harrison Tanzola, and Kristina Tummino for assistance with various aspects of the study.

Conflicts of interest: none.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.concog.2018.12.001.

References


McFarland, C. P., Primosch, M., Maxson, C. M., & Stewart, B. T. (2017). Enhancing memory and imagination improves problem solving among individuals with...


