Preparing for what might happen: An episodic specificity induction impacts the generation of alternative future events

Helen G. Jing\textsuperscript{a,b,∗}, Kevin P. Madore\textsuperscript{b}, Daniel L. Schacter\textsuperscript{a}

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

A R T I C L E   I N F O
Keywords:
Episodic future simulation
Prospection
Episodic specificity induction
Alternative future events
Emotion
Debiasing

A B S T R A C T
A critical adaptive feature of future thinking involves the ability to generate alternative versions of possible future events. However, little is known about the nature of the processes that support this ability. Here we examined whether an episodic specificity induction—brief training in recollecting details of a recent experience that selectively impacts tasks that draw on episodic retrieval—(1) boosts alternative event generation and (2) changes one’s initial perceptions of negative future events. In Experiment 1, an episodic specificity induction significantly increased the number of alternative positive outcomes that participants generated to a series of standardized negative events, compared with a control induction not focused on episodic specificity. We also observed larger decreases in the perceived plausibility and negativity of the original events in the specificity condition, where participants generated more alternative outcomes, relative to the control condition. In Experiment 2, we replicated and extended these findings using a series of personalized negative events. Our findings support the idea that episodic memory processes are involved in generating alternative outcomes to anticipated future events, and that boosting the number of alternative outcomes is related to subsequent changes in the perceived plausibility and valence of the original events, which may have implications for psychological well-being.

1. Introduction
In recent years, there has been a surge in research focusing on episodic simulation or the construction of a detailed representation of a hypothetical personal future experience (Schacter, Addis, & Buckner, 2008), as well as other types of prospective, future-oriented thought (Seligman, Railton, Baumeister, & Srípada, 2013; Szpunar, Spreng, & Schacter, 2014). According to the constructive episodic simulation hypothesis (Schacter & Addis, 2007), having a constructive, flexible episodic memory system allows individuals to imagine or simulate future scenarios by drawing on one’s past experiences. This hypothesis has since been supported by extensive research highlighting the striking cognitive and neural similarities between episodic memory and episodic future simulation. For example, populations with deficits in recalling past events face similar challenges when asked to imagine future experiences, and there are significant overlaps in both the phenomenological quality and neural regions recruited when recalling the past and simulating the future (for recent reviews, see Benoit & Schacter, 2015; Schacter et al., 2012; Szpunar, 2010). Recent work has explored not only the mechanisms that support episodic future simulation, but also the functions that future thinking may serve (for review, see Schacter, Benoit, & Szpunar, 2017). The current paper aims to examine both the episodic memory mechanisms underlying our ability to engage in a specific type of future thinking—generating alternative outcomes to future events—and adaptive ways in which simulating alternative future events subsequently impacts our original expectations and perceptions of the future.

David Ingvar (1979, p. 21) previously theorized about the adaptive functions of future thinking, stating, “On the basis of previous experiences, represented in memories, the brain—one’s mind—is automatically busy with extrapolation of future events and, as it appears, constructing alternative hypothetical behavior patterns in order to be ready for what may happen” (see also Schacter, 2012; Suddendorf & Corballis, 1997, 2007). Notably, the hypothesized preparatory function of future thinking has been demonstrated in the context of planning and problem solving (e.g., Arnold, Iaria, & Ekstrom, 2016; Gerlach, Spreng, Gilmore, & Schacter, 2011; Spreng, Stevens, Chamberlain, Gilmore, & Schacter, 2010), prospective memory (e.g., Terrett et al., 2016), decision-making (e.g., Benoit, Gilbert, & Burgess, 2011; Peters & Büchel, 2010), and emotion regulation (e.g., Taylor,
PHAM, Rivkin, & ARMOR, 1998; for review, see Schacter, 2012). However, despite Ingvär’s initial emphasis on the importance of constructing alternative behavior patterns, there remains very little investigation of the actual simulation of alternative versions of future events.

Although there is a large literature on counterfactual thinking, or generating alternative outcomes and consequences for autobiographical past events (for recent reviews, see Byrne, 2016; Epstude & Roese, 2008; for related work on mental models of alternate possibilities, see Johnson-Laird, 2001), most research on future event simulation focuses on the construction of complex scenes and action sequences within the specific context of a single event, without considering possible alternative outcomes for such an event. There does exist a smaller body of research on prefactual thinking, a type of conditional future thought concerning alternative future outcomes that may occur with some degree of certainty (Byrne & Egan, 2004; Epstude, Scholl, & Roese, 2016; Petrocchi, Seta, & Seta, 2012). Prefactuals often take the form of, “If I take action X, then it may lead to outcome Y”; for example, “If I study for 5 hours, then I may get a better grade on my exam, or “If I only study for 1 hour, then I may do poorly” (for related work on implementation intentions, see Gollwitzer, 1999; Gollwitzer & Sheeran, 2006). However, there are many instances when individuals tend to harbor expectations for an emotional future event outcome (e.g., “I will do poorly on the exam next week”), without necessarily considering conditional features of the event in a “what-if” or “if-then” manner (Norem & Cantor, 1986; Oettingen & Mayer, 2002; Wilson & Gilbert, 2005). These predicted emotional outcomes can subsequently color one’s actual experience of the event (Showers, 1992; Wilson, Lisle, Kraft, & Wetzel, 1989). Thus, there remains much to be explored regarding the consideration of alternative outcomes to an expected future experience.

One potential benefit of generating alternative future outcomes relates to emotion regulation. Engaging in constructive, positive thoughts towards negative future events is known to be beneficial in a variety of ways, such as decreasing negative emotions and worry (Brown, MacLeod, Tata, & Goddard, 2002; Jing, Madore, & Schacter, 2016; MacLeod, 2017) and increasing engagement in active strategies to cope with a stressor (Pham & Taylor, 1999; Rivkin & Taylor, 1999; for review, see Taylor et al., 1998). While episodic simulation is often adaptive in these contexts, there are situations in which simulation can be disruptive to one’s psychological well-being. For example, in emotional disorders such as anxiety or depression, individuals may show greater anticipation of negative future experiences (e.g., MacLeod & Byrne, 1996) and harbor excessive worry about the future (e.g., Borkovec, Ray, & Stöber, 1998; see also Bulley, Henry, & Suddendorf, 2017; Miloyan, Pachana, & Suddendorf, 2014). Negative thoughts towards the future can be maladaptive if one repeatedly engages in those thoughts, particularly given that repetitive future thinking has been linked to increased estimates in the perceived plausibility of an event’s occurrence (Szpunar & Schacter, 2013; Wu, Szpunar, Godovich, Schacter, & Hofmann, 2015) and increased accessibility of negative event outcomes (Byrne & MacLeod, 1997; MacLeod, Tata, Kentish, Carroll, & Hunter, 1997). Thus, it is critical to interrupt the cycle of repetitive thinking and reevaluate the perceived plausibility of anticipated negative future experiences, and one way to do so may be to consider alternative outcomes to negative future events.

Bentz, Williamson, and Franks (2004) proposed that using techniques to generate alternative, more positive outcomes may serve as a “debiasing” strategy to reduce pessimistic likelihood judgments of the future. If anxious individuals operate under the availability heuristic (Tversky & Kahneman, 1973) and judge negative outcomes to be highly plausible because they are highly accessible, then activating positive outcomes should increase the ease by which those outcomes may come to mind, suggesting that the anticipated event may not be as predictable as previously believed. Thus, considering positive alternatives may subsequently decrease the perceived likelihood that the original negative event will occur. Bentz et al. (2004) used a Consider-An-Alternative debiasing strategy (cf. Hirt & Markman, 1995) and found that asking both highly trait-anxious and non-anxious individuals to generate three alternative positive outcomes for a variety of negative scenarios significantly reduced participants’ ratings of event plausibility (see also Hirt, Kardes, & Markman, 2004; Raune, MacLeod, & Holmes, 2005). However, Bentz et al. (2009) also demonstrated that generating more alternative outcomes (e.g., five positive outcomes) incurred no additional benefit in debiasing judgments when compared to generating fewer alternative outcomes (e.g., two positive outcomes). One possible explanation is that task difficulty increases when participants are asked to generate a larger number of alternative outcomes, presumably decreasing the realistic and positive quality of the generated alternatives.

While Bentz et al. (2004, 2009) demonstrated that considering positive alternative event outcomes, rather than fixating on a single future experience, is beneficial for debiasing pessimistic judgments of the future, little is known about the processes that support alternative event simulation, and fundamental issues remain to be addressed. The main goal of the present studies is to enhance our basic understanding of how people generate and simulate alternative outcomes to future events by addressing two issues that emerge from the earlier work of Bentz et al. (2004, 2009).

First, utilizing a tool that can boost the number of positive alternative event outcomes as well as boost the ease of generating these outcomes may help to determine whether the ease of alternative outcome generation is truly a limiting factor in the demonstrated effects of the Consider-An-Alternative debiasing intervention. One such tool is an episodic specificity induction, a brief training in recollecting details of a recent experience (Madore, Gaesser, & Schacter, 2014). This procedure is based on the Cognitive Interview (CI; Fisher & Geiselman, 1992; Memon, Meissner, & Fraser, 2010), a well-established method that increases recall of episodic detail in eyewitnesses. Prior work on the specificity induction has shown that encouraging participants to focus on specific details of a past experience (e.g., of a short video they just watched or of a past autobiographical memory) selectively biases participants to focus on specific event details during subsequent tasks that are dependent on episodic memory; consequently, individuals tend to construct more detailed mental scenes or events after a specificity induction than a control induction (for further discussion, see Schacter & Madore, 2016).

Apart from increasing the level of detail when recalling past and imagining future experiences (Madore et al., 2014), other known effects of the specificity induction on subsequent tasks include boosting the number of steps generated during problem solving (Jing et al., 2016; Madore & Schacter, 2014), increasing the number of creative solutions generated during divergent thinking tasks (Madore, Addis, & Schacter, 2015; Madore, Jing, & Schacter, 2016), and boosting the level of concreteness and detail during episodic reappraisal (Jing et al., 2016). Critically, whereas the specificity induction impacts performance on subsequent tasks that are thought to be dependent on episodic memory, it has no detectable impact on the performance of tasks that are thought to rely on primarily semantic retrieval or non-episodic narrative processing, such as describing a picture (Madore et al., 2014), generating word definitions (Madore & Schacter, 2016), or generating object associations and semantic solution words (Madore et al., 2015). We suggest that the generation and simulation of alternative event outcomes depends critically upon episodic memory, in a similar manner as other types of future event simulation. Accordingly, we predict that the specificity induction will increase the number of positive alternatives that participants generate, relative to a control induction. In addition, the specificity induction should boost the ease with which individuals generate these positive alternative outcomes, allowing us to examine whether such a boost may be linked to further debiasing of pessimistic future judgments.

Second, while event plausibility is a construct that influences one’s psychological well-being towards the event in question (Bentz et al.,
perceived event valence is also important. Research on affective forecasting has shown that individuals are quite inaccurate at estimating their emotional reactions to future experiences, often displaying an impact bias, whereby they overestimate the intensity of their emotional reactions, particularly for negative events (Andrade & Van Boven, 2010; Gilbert & Wilson, 2009; Wilson & Gilbert, 2005). Some proposed causes of these errors include underestimating the influence that other events or other aspects of the event in question may have on one’s thoughts and feelings, as well as misjudging how well one might emotionally adapt to the situation (Gilbert & Wilson, 2009; Wilson & Gilbert, 2005). Here, we suggest that generating alternative positive outcomes for a negative event might reduce the perceived negativity of the original event. Considering alternative positive outcomes grants access to additional information about ways in which the event may unfold, including previously ignored details about other situational and emotional factors (Gilbert & Wilson, 2009; Taylor et al., 1998). This additional information, in turn, may lead one to realize that the event might not be as negative as initially perceived, and thus encourage a subsequent adjustment of negative expectations.

In the current experiments, we tested the hypothesis that manipulating the number of alternative future outcomes that individuals generate would influence subsequent measures of perceived event valence and plausibility. Based on previous findings, we predicted that the episodic specificity induction, relative to a control induction, should (1) increase the number of relevant positive alternative scenarios that participants generate, (2) boost the perceived ease of generating alternative event scenarios, and (3) decrease the perceived negativity and plausibility of the original event.

2. Experiment 1

2.1. Method

2.1.1. Participants

A total of 33 healthy undergraduate students (ages 18–25, M = 20.93 years, 22 female) were recruited from Harvard University. All participants had normal vision and no history of neurological or psychological impairment, and were paid or received course credit for their participation. A total of 7 participants were excluded due to failure to complete the experiment (4 participants), noncompliance (2 participants), or inability to perform the experimental tasks (1 participant), leaving 26 participants in the final sample. Before the study was run, we performed a power analysis (G’Power 3; see Faul, Erdfelder, Lang, & Buchner, 2007) to determine that a sample size of at least 23 usable participants was necessary to observe a medium-sized effect of the induction (power > 0.80, α = 0.05, two-tailed, for a within-subjects design, d = 0.62), which has also been the case in prior induction studies (e.g., Jing et al., 2016; Madore et al., 2014, 2015). Given scheduling constraints with multiple sessions, data collection was stopped once approximately enough usable participants had been run to reach this number.

2.1.2. Equipment

All experimental sessions were executed using Qualtrics on an Apple desktop computer. Participants viewed the induction videos using Quicktime media player, and verbal responses during the induction questioning procedures were recorded using an audio recorder.

2.1.3. Experimental procedure

The experiment as a whole lasted 4 hours across 2 separate sessions that took place 4–7 days apart (M = 5.19 days). Both experimental sessions lasted 2 hours and were very similar in structure, consisting of several different phases: event simulation, induction phase (specificity or control induction), alternative event generation, and finally an alternative event rating phase.

2.1.3.1. Event stimuli. Across both experimental sessions, participants viewed a total of 12 standardized negative events that were specific, concrete, and had tangible outcomes that could plausibly occur within the next several years. These 12 event stimuli were scenarios that 30 separate participants had most frequently generated during a pilot study, during which they were asked to list a series of plausible and familiar negative future events that spanned across a variety of domains such as academics, health, career, relationships, and finances. The 12 events were separated into two lists of 6 events each, and presentation of the two lists during the two experimental sessions was counterbalanced across participants (see Supplemental Materials for event lists).

2.1.3.2. Event simulation phase. In each session, participants were presented with 6 standardized negative events. For each separate event, participants were first provided with the event title (e.g., “I receive a bad grade on my exam”) and were instructed to imagine experiencing the concrete future scenario specified by the title for 2 min. They were told that each event should take place in a specific location within a time frame of several hours, and were encouraged to elaborate upon their negative thoughts and feelings in as much detail as possible. Afterwards, participants rated each event for valence, the perceived plausibility of occurrence, and the similarity of the future event to something they had previously experienced on a scale of 1–9. Participants first completed a practice trial with the experimenter to ensure they understood all instructions.

2.1.3.3. Induction phase. The specificity and control inductions were administered in the same manner as in our previous studies (e.g., Jing et al., 2016; Madore et al., 2014). Following event simulation, participants watched a short video of two adults performing routine activities in a kitchen; two different videos were used between induction conditions and the order of videos was counterbalanced across subjects. Following the video, participants completed a 2-min math filler task comprised of addition and subtraction questions. Next, participants received questions about the video in the form of either an episodic specificity induction or an impressions control induction; only one induction was administered per session and the order of inductions was counterbalanced across subjects. In the episodic specificity induction, participants were given mental imagery probes asking them to recall specific details about the setting, people and actions in the video, with follow-up probes asking them to elaborate more on the details they had mentioned. In the impressions control induction, participants were asked questions targeting their general impressions and thoughts about the video, which allowed them to talk more generally about the video without requiring them to retrieve specific episodic details (see Supplemental Materials for induction scripts). Contrasting the two induction conditions allowed us to assess the effect of a more specific retrieval orientation (as induced by the specificity induction) on participants’ subsequent abilities to generate alternative scenarios, as well as potential effects of a boost in alternative event generation on the perceived plausibility and valence of the original events.

2.1.3.4. Alternative event generation phase. After receiving one induction, participants viewed the titles of each of the 6 negative events they simulated in the first phase. For each event, participants were given 5 min to generate and type out as many alternative event lists).
portion of the final grade), or changing the imagined interactions they had with others in the original event (e.g., finding out about the grade over email instead of directly hearing from the professor). Critically, participants were instructed to generate alternative scenarios that were more positive than the original negative event. In addition, they were instructed to keep constant the core elements of the original event; for example, if the original event concerned a specific academic class, in all alternative event scenarios they should continue thinking about the same class and refrain from thinking about alternative scenarios in other classes.

After generating as many alternative event scenarios as possible for 5 min, participants were then asked to make three ratings on a scale of 1–9: (1) the ease or difficulty of generating alternative scenarios for each event, (2) the perceived valence of the original event, and (3) the perceived plausibility of the original event. Once again, participants first completed a practice trial with the experimenter to ensure they understood all instructions.

2.1.3.5. Alternative event rating phase. Next, participants were presented with each alternative scenario that they listed in the previous phase (i.e., the exact sentence or phrase they typed during alternative event generation), and were asked to rate each scenario on five different dimensions on a scale of 1–9: (1) valence, (2) plausibility, (3) similarity of the alternative scenario to a past experience, (4) similarity of the alternative scenario to something they had previously thought about (without actual experience), and (5) how realistic they perceived the alternative scenario to be.

2.2. Coding

Three raters were trained to score responses from the 5-min alternative event generation task. Responses were scored as either a relevant or irrelevant alternative event scenario. A relevant alternative event is one that keeps constant the core elements of the original negative scenario while changing other aspects of the event, such as the outcome, emotional content, perceived consequences, and other imagined interactions from the original event (for more detail, see the alternative event generation phase description in Section 2.1.3). An irrelevant alternative event is one that either (1) changes a core component of the original event (e.g., describing an alternative scenario concerning a history exam when the original event concerned a math exam), (2) does not pertain to the original scenario in a coherent or logical manner (e.g., for an original event where one does poorly on an exam, generating an alternative scenario where one visits a museum), or (3) does not actually change elements of the original scenario. We drew the distinction between relevant and irrelevant alternative outcomes to match the scoring procedures in prior work examining the effect of the specificity induction on generating steps to solve a problem (e.g., “relevant steps” vs. “no steps”; Jing et al., 2016; Madore & Schacter, 2014) and on details in memory and imagination tasks (e.g., “internal”, episodic event details that are relevant to the task vs. “external”, semantic details or other details deemed irrelevant to the task; Madore et al., 2014).

All raters were blind to the condition (control, specificity) of participant responses. The three raters separately scored 20 sample participant responses to assess inter-rater reliability, and high inter-rater reliability was obtained for both types of alternatives (standardized Cronbach’s $\alpha = 0.99$ for relevant alternatives and 0.92 for irrelevant alternatives). The remainder of responses was scored separately by one of the three raters. Rater 1 scored 38% of participant responses, rater 2 scored 31% of participant responses, and rater 3 scored 31% of participant responses.

2.3. Results

We conducted a series of repeated-measures analyses of variance (ANOVs) to test the hypotheses, which involved within-subjects factors of Induction condition (control vs. specificity), Alternative event type (relevant vs. irrelevant), and Time of rating (initial ratings prior to generating alternative events vs. ratings after generating alternative events). Both main effects and interactions were tested for each of the variables, but the interactions trumped the main effects and explicitly addressed our hypotheses. The counterbalanced order of induction did not have a significant effect on the following analyses. We subsequently conducted a series of linear multilevel models that examine the relationship between the variables of interest on an event trial-by-trial basis, rather than aggregating data into participant means.

2.3.1. Initial event ratings

There were no significant differences in initial ratings of event valence between the two lists collapsed across both control and specificity conditions, $M_{\text{difference}} = −0.08, SE = 0.20, t(24) = −0.43, p > 0.250$, 95% CI = $[−0.49, 0.32], d = 0.17$. We also did not find significant differences in initial ratings of event plausibility between lists, $M_{\text{difference}} = −0.09, SE = 0.39, t(24) = −0.23, p > 0.250$, 95% CI = $[−0.89, 0.71], d = 0.09$.

2.3.2. Induction effects on alternative event generation

We first examined how the specificity induction affected the number of alternative event scenarios that participants generated. Participants spent slightly longer discussing the video in the specificity induction ($M = 4.74$ min, $SD = 0.87$) than in the control induction ($M = 4.02$ min, $SD = 0.60$), $M_{\text{difference}} = 0.72, SE = 0.10, t(25) = 6.89, p < 0.001$, 95% CI = $[0.50, 0.93], d = 1.35$. However, including the difference score for induction duration as a covariate in the following repeated-measures ANOVAs did not significantly affect any results.

We conducted a 2 (Induction: control vs. specificity) × 2 (Alternative event type: relevant vs. irrelevant) repeated-measures ANOVA. Results revealed significant main effects of Induction, $F(1,25) = 7.15, p = 0.013, \eta^2_p = 0.22$, and Alternative event type, $F(1,25) = 195.84, p < 0.001, \eta^2_p = 0.89$, but critically, we found a significant interaction of Induction × Alternative event type, $F(1,25) = 37.47, p < 0.001, \eta^2_p = 0.60$. Two-tailed post hoc t-tests showed that participants generated significantly more relevant alternative event scenarios in the specificity condition compared to the control condition, $M_{\text{difference}} = 1.05, SE = 0.20, t(25) = 5.20, p < 0.001$, 95% CI = $[0.63, 1.46], d = 1.02$, whereas the induction had no effect on the number of irrelevant alternative event scenarios that were generated, $M_{\text{difference}} = −0.13, SE = 0.19, t(25) = −0.68, p > 0.250$, 95% CI = $[−0.52, 0.26], d = 0.13$ (Fig. 1). Thus, the specificity induction selectively boosted the number of relevant alternative event scenarios produced.

Fig. 1. Experiment 1 mean induction effects on the number of relevant and irrelevant alternative event scenarios generated in the control and specificity conditions, where the y-axis represents the mean number of alternative events per trial. Error bars represent one standard error of the mean.

121
that participants generated (see Supplemental Table 1 for mean values). For effects of the specificity induction on details contained within the generated alternative events (i.e., scored using the Autobiographical Interview; Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002), see Supplemental Materials.

2.3.3. Perceived difficulty and realistic quality of alternative events

Participants rated the process of generating alternative event scenarios in the specificity condition as less difficult than in the control condition, $M_{\text{difference}} = -0.59$, $SE = 0.21$, $t(25) = -2.77$, $p = 0.011$, 95% CI = [-1.03, -0.15], $d = 0.54$. Furthermore, there was no significant difference in the average ratings of how realistic participants perceived the alternative events to be in the control (i.e., fewer alternative events) and specificity (i.e., more alternative events) conditions, $M_{\text{difference}} = -0.18$, $SE = 0.16$, $t(25) = -1.14$, $p > 0.250$, 95% CI = [-0.51, 0.15], $d = 0.22$ (see Supplemental Table 1 for mean values).

2.3.4. Induction effects on valence and plausibility ratings

Because we found that the specificity induction boosted the number of alternative events that participants generated, next we related this boost to changes in perceptions of the original negative events by comparing mean ratings of event valence and plausibility immediately before and after generating alternative events in both the control and specificity conditions. We first report analyses that examine the contrast between rating changes in the control and specificity conditions through a series of 2 (Induction: control vs. specificity) $\times$ 2 (Time of rating: before vs. after generating alternatives) repeated-measures ANOVAs. We then report linear multilevel models that examine the relationship between the variables of interest on a trial-by-trial basis.

For ratings of perceived original event valence, there was no main effect of Induction, $F(1,25) = 1.52$, $p = 0.229$, $\eta^2_p = 0.06$, and a significant main effect of Time of rating, $F(1,25) = 53.70$, $p < 0.001$, $\eta^2_p = 0.68$. Most importantly, we found a significant interaction of Induction $\times$ Time of rating, $F(1,25) = 7.76$, $p = 0.010$, $\eta^2_p = 0.24$. There was a significant decrease in the perceived negativity of the imagined events from before to after alternative event generation in both the control and specificity conditions (see Supplemental Table 1 for mean values), but critically, there was a significantly larger decrease in ratings of negative valence in the specificity condition than in the control condition, $M_{\text{difference}} = -0.45$, $SE = 0.16$, $t(25) = -2.79$, $p = 0.010$, 95% CI = [-0.79, -0.12], $d = 0.55$ (Fig. 2A). Furthermore, across all participants there was a trending negative relationship between the average boost in the number of alternative event scenarios that participants generated (i.e., the difference in the number of alternatives generated in the specificity condition vs. the control condition) and the average decrease in perceived negativity of the original events (i.e., the difference in the valence rating before vs. after generating alternative outcomes) between the specificity and control induction conditions, $r(24) = -0.368$, $p = 0.064$ (Fig. 2B).

Next, we ran a linear multilevel model to further examine the relationship between the number of alternatives generated and the perceived change in negativity of the original events by induction condition (control vs. specificity), with events as a level one predictor and participants as a level two predictor. This analysis allowed us to examine whether the number of alternative events generated predicted the change in negativity of the original event on a trial-by-trial basis. Both the induction condition and the number of alternative events were treated as fixed-effect predictor variables, and the interaction between the number of alternative events and participants was treated as a random effect. The outcome variable of interest was the change in perceived negativity of the original event. Induction condition was not a significant predictor, but the number of alternative events that participants generated significantly predicted the change in negativity of the original event, $B = -0.19$, $t(197.43) = -3.57$, $p < 0.001$. That is, the more alternatives that participants generated per event trial, the greater the observed decrease in perceived negativity towards the original event.

For ratings of perceived plausibility of the original negative events, there was no main effect of Induction, $F(1,25) = 0.03$, $p > 0.250$, $\eta^2_p = 0.001$, and a significant main effect of Time of rating, $F(1,25) = 35.97$, $p < 0.001$, $\eta^2_p = 0.59$. Critically, we found a significant interaction of Induction $\times$ Time of rating, $F(1,25) = 4.60$, $p = 0.042$, $\eta^2_p = 0.16$. We found a significant decrease in the perceived plausibility of the imagined events from before to after alternative event generation in both the control and specificity conditions (see Supplemental Table 1 for mean values), but there was a larger decrease in ratings of perceived event plausibility in the specificity condition than in the control condition, $M_{\text{difference}} = -0.53$, $SE = 0.25$, $t(25) = -2.15$, $p = 0.042$, 95% CI = [-1.04, -0.02], $d = 0.42$ (Fig. 2C). We also observed a significant negative relationship between the average boost in the number of alternative event scenarios and the average decrease in perceived plausibility of the original events between the specificity and control induction conditions across all participants, $r(24) = -0.420$, $p = 0.033$ (Fig. 2D).

Once again, we ran a linear multilevel model to examine the relationship between the number of alternatives generated and the perceived change in plausibility of the original events by induction condition, with events as a level one predictor and participants as a level two predictor. Fixed-effect predictors included induction condition and the number of alternative events, the interaction between the number of alternative events and participants was treated as a random effect, and the outcome variable was the change in perceived plausibility of the original event. We found that the number of alternative events that participants generated significantly predicted the change in plausibility of the original event, $B = -0.30$, $t(133.39) = -5.05$, $p < 0.001$, such that the more alternatives that participants generated per event trial, the greater the reported decrease in perceived plausibility towards the original event.

Overall, these results suggest that the boost in the number of generated alternative event scenarios via the specificity induction is related to larger reductions in the perceived negativity and perceived plausibility of the original negative future events. On a trial-by-trial level, the number of alternative events was also negatively related to subsequent changes in perceived negativity and plausibility of the original events.

2.4. Experiment 1 discussion

Overall, the results of Experiment 1 support the hypothesis that inducing a more specific and detailed retrieval orientation via an episodic specificity induction boosts the number of alternative event scenarios that participants generated. Furthermore, participants rated the process of generating alternative events as less difficult in the specificity condition than in the control condition, and the alternative events were rated as similarly plausible between conditions.

Importantly, we observed larger decreases in the perceived negativity and plausibility of the original events in the specificity condition, where participants generated more alternative outcomes, relative to the control condition. In contrast to the experiments by Bentz et al. (2004, 2009), by using the specificity induction we were able to further decrease the perceived negativity and plausibility of the original negative events without observing a reported increase in task difficulty. In accordance with our hypotheses, these results suggest that thinking about positive alternative outcomes to negative events may alter one’s initial perception of those events. In Experiment 2, we aimed to replicate and extend our findings from Experiment 1 to a series of personalized (rather than standardized) negative future experiences.

3. Experiment 2

In Experiment 2, we aimed to examine the effect of an episodic specificity induction on the generation of alternative outcomes for a series of personalized negative events. Rather than using a set of
standardized event cues (as in Experiment 1), participants were directly asked to list a series of anticipated negative events that may occur in their own personal future. Overall, the methods used in Experiment 2 are very similar to those of Experiment 1, with differences highlighted below.

3.1. Method

3.1.1. Participants

A total of 36 healthy undergraduates were recruited from Harvard University and Boston University (ages 18–25, M = 21.42 years, 19 female). A total of 7 participants were excluded due to noncompliance (1 participant) or incompletion of the experiment (6 participants), leaving 29 participants in the final analysis. To keep the sample size in Experiment 2 comparable to that of Experiment 1, we stopped data collection after reaching a similar number of useable participants.

3.1.2. Experimental procedure

On average, session 2 took place 5.86 days after session 1. As in Experiment 1, both experimental sessions lasted 2 hours and were very similar in structure, consisting of several different phases: event simulation (i.e., imagining 6 negative events for 2 min each), induction phase (specificity or control induction), alternative event generation (i.e., generating alternative event outcomes for 5 min), and finally an alternative event rating phase (i.e., rating each alternative event for valence, plausibility, and novelty). The only procedural change from Experiment 1 took place during the event simulation phase, which we describe below.

3.1.2.1. Event generation and simulation phase

In each session, participants first provided 6 negative events that were specific, concrete, highly familiar, and had tangible outcomes that could plausibly occur within the next several years. Across both experimental sessions, participants generated a total of 12 negative events that concerned topics relating to academics, health, career, relationships, and finances. For each separate event, they were asked to provide a brief title and then imagine a concrete future scenario in which they were experiencing the event in as much detail as possible for 2 min. Afterwards, participants rated each original event for valence, plausibility, and similarity to previous experience on a 1–9 scale.

3.1.3. Coding

Two raters (raters 1 and 3 from Experiment 1) were trained to score responses from the 5-min alternative event generation task as either relevant or irrelevant alternative event scenarios. Both raters were blind to the induction condition of participant responses. High inter-rater reliability was obtained for both types of alternatives (standardized Cronbach’s α = 0.99 for relevant alternatives and 0.90 for irrelevant alternatives). The two raters scored 66% and 34% of participant responses, respectively.

3.2. Results

As in Experiment 1, we conducted repeated-measures ANOVAs to test the hypotheses, which involved within-subjects factors of Induction
condition (control vs. specificity), Alternative event type (relevant vs. irrelevant), and Time of rating (initial ratings prior to generating alternative events vs. ratings after generating alternative events). The counterbalanced order of induction did not have a significant effect on the analyses reported below. We also report a series of linear multilevel models that examine the relationship between the variables of interest on an event trial-by-trial basis.

3.2.1. Induction effects on alternative event generation

Participants spent slightly longer discussing the video in the specificity induction (M = 4.73 min, SD = 1.24) than in the control induction (M = 3.95 min, SD = 0.68), $t(28) = 2.90$, $p = 0.009$, 95% CI = [0.37, 1.19], $d = 0.72$. However, including the difference score for induction duration as a covariate in the following repeated-measures ANOVAs did not significantly affect any results.

We first conducted a 2 (Induction: control vs. specificity) × 2 (Alternative event type: relevant vs. irrelevant) repeated-measures ANOVA, and found significant main effects of Induction, $F(1,28) = 5.13$, $p = 0.031$, $\eta^2_p = 0.16$, and Alternative event type, $F(1,28) = 458.65$, $p < 0.001$, $\eta^2_p = 0.94$. As in Experiment 1, there was a significant interaction of Induction × Alternative event type, $F(1,28) = 22.66$, $p < 0.001$, $\eta^2_p = 0.45$. Participants generated significantly more relevant alternative event scenarios in the specificity condition compared to the control condition, $M_{\text{difference}} = 1.29$, $SE = 0.25$, $t(28) = 5.14$, $p < 0.001$, 95% CI = [0.78, 1.80], $d = 0.95$. There was also a trend towards generating fewer irrelevant alternative event scenarios in the specificity condition than in the control condition, $M_{\text{difference}} = -0.52$, $SE = 0.26$, $t(28) = -2.00$, $p = 0.056$, 95% CI = [−1.05, 0.01], $d = 0.37$, although this difference did not reach full significance (Fig. 3). Thus, the specificity induction significantly boosted the number of relevant alternative event scenarios that participants generated (see Supplemental Table 2 for mean values).

3.2.2. Perceived difficulty and realistic quality of alternative events

As in Experiment 1, participants rated the process of generating alternative event scenarios in the specificity condition as less difficult than in the control condition, $M_{\text{difference}} = -0.90$, $SE = 0.29$, $t(28) = -3.11$, $p = 0.004$, 95% CI = [−1.49, −0.31], $d = 0.58$. Furthermore, there was no significant difference in the average ratings of how realistic participants perceived the alternative events to be in the control and specificity conditions, $M_{\text{difference}} = -0.15$, $SE = 0.13$, $t(28) = -1.17$, $p > 0.250$, 95% CI = [−0.41, 0.11], $d = 0.22$ (see Supplemental Table 2 for mean values).

3.2.3. Induction effects on valence and plausibility ratings

Next, we contrasted the mean ratings of event valence and perceived plausibility between the control and specificity conditions through a series of 2 (Induction: control vs. specificity) × 2 (Time of rating: before vs. after generating alternatives) repeated-measures ANOVAs. For ratings of perceived original event valence, results revealed no main effect of Induction, $F(1,28) = 2.39$, $p = 0.133$, $\eta^2_p = 0.07$, and a significant main effect of Time of rating, $F(1,28) = 34.97$, $p < 0.001$, $\eta^2_p = 0.56$. Most importantly, we found a significant interaction of Induction × Time of rating, $F(1,28) = 7.20$, $p = 0.012$, $\eta^2_p = 0.21$. There was a significant decrease in the perceived negativity of the imagined events from before to after alternative event generation in both the control and specificity conditions (see Supplemental Table 2 for mean values), but critically, there was a significantly larger decrease in ratings of negative valence in the specificity condition than in the control condition, $M_{\text{difference}} = -0.46$, $SE = 0.17$, $t(28) = -2.68$, $p = 0.012$, 95% CI = [−0.80, −0.11], $d = 0.50$ (Fig. 4A). We also observed a significant negative relationship between the boost in the number of alternative event scenarios (i.e., the difference in the number of alternatives generated in the specificity vs. control condition) and the decrease in perceived negativity of the original events (i.e., the difference in the valence rating before vs. after generating alternative outcomes) between the specificity and control induction conditions, $r(27) = -0.381$, $p = 0.041$ (Fig. 4B).

As in Experiment 1, we ran a linear multilevel model to further examine the relationship between the number of alternatives generated and the perceived change in negativity of the original events by induction condition (control vs. specificity), with events as a level one predictor and participants as a level two predictor. This analysis allowed us to examine whether the number of alternative events generated predicted the change in negativity of the original event on a trial-by-trial basis. Induction condition and the number of alternative events were treated as fixed-effect predictor variables, and the interaction between the number of alternative events and participants was treated as a random effect. The outcome variable of interest was the change in perceived negativity of the original event. Induction condition was not a significant predictor, but the number of alternative events that participants generated significantly predicted the change in negativity of the original event, $B = -0.21$, $t(284.19) = -3.95$, $p < 0.001$. That is, the more alternatives that participants generated per event trial, the greater the observed decrease in perceived negativity towards the original event.

For ratings of perceived plausibility of the original negative events, we found no main effect of Induction, $F(1,28) = 1.22$, $p > 0.250$, $\eta^2_p = 0.04$, and a significant main effect of Time of rating, $F(1,28) = 60.16$, $p < 0.001$, $\eta^2_p = 0.68$. We observed a significant interaction of Induction × Time of rating, $F(1,28) = 9.72$, $p = 0.004$, $\eta^2_p = 0.26$. There was a significant decrease in the perceived plausibility of the imagined events from before to after alternative event generation in both the control and specificity conditions (see Supplemental Table 2 for mean values), but there was a larger decrease in ratings of perceived event plausibility in the specificity condition than in the control condition, $M_{\text{difference}} = -0.44$, $SE = 0.14$, $t(28) = -3.12$, $p = 0.004$, 95% CI = [−0.73, −0.15], $d = 0.58$ (Fig. 4C). There was also a weak negative relationship between the boost in the number of alternative event scenarios and the decrease in perceived plausibility of the original events between the specificity and control induction conditions, $r(27) = -0.318$, $p = 0.093$ (Fig. 4D).

We also ran a linear multilevel model to examine the relationship between the number of alternatives generated and the perceived change in plausibility of the original events on a trial-by-trial basis. Induction condition and the number of alternative events were treated as fixed-effect predictor variables, the interaction between the number of alternative events and participants was treated as a random effect, and the outcome variable of interest was the change in perceived plausibility of the original event. Induction condition was not a significant
predictor, but the number of alternative events that participants generated significantly predicted the change in plausibility of the original events, $B = -0.12$, $t(254.44) = -2.32$, $p = 0.021$. Thus, the more alternatives that participants generated per event trial, the greater the observed decrease in perceived plausibility towards the original event.

As in Experiment 1, these results suggest that the boost in the number of generated alternative event scenarios via the specificity induction is related to larger reductions in the perceived negativity and plausibility of the original events.

3.3. Experiment 2 discussion

The results of Experiment 2 replicate the results of Experiment 1 even when using personalized rather than standardized negative future scenarios. These data once again support the idea that using an episodic specificity induction boosted the perceived ease of simulation and increased the number of alternative event scenarios that participants generated, while leading to larger decreases in the perceived negativity and plausibility of the original events in the specificity condition relative to the control condition. On a trial-by-trial level, the number of alternative events was also negatively related to subsequent changes in perceived negativity and plausibility of the original events. Thus, thinking about alternative positive outcomes to negative events can have beneficial effects on one’s perception of those events.

B. Boost in Alternatives vs. Negativity Change

Fig. 4. Experiment 2 effects of generating alternative event scenarios on original event valence and plausibility: (A) mean induction effects on changes in valence ratings before and after alternative event generation, where the y-axis represents the mean rating per trial and error bars represent one standard error of the mean, (B) the relationship between the average boost in alternative event generation and the average decrease in original event negativity between the control and specificity conditions across all participants, (C) mean induction effects on changes in plausibility ratings before and after alternative event generation, where the y-axis represents the mean rating per trial and error bars represent one standard error of the mean, and (D) the relationship between the average boost in alternative event generation and the average decrease in original event plausibility between the control and specificity conditions across all participants.

4. General discussion

Overall, the data from two experiments support the hypothesis that increasing the number of different ways in which individuals consider how future events may unfold may subsequently alter one’s subjective perception of those events. First, using an episodic specificity induction increased the number of alternative events that participants generated. While the specificity induction has previously been shown to enhance performance on a variety of tasks thought to be dependent upon episodic memory (for review, see Schacter & Madore, 2016), we demonstrate for the first time that the induction also affects the process of generating alternative outcomes to a series of anticipated negative future events. It is important to note that we do not wish to imply that episodic memory is the only form of memory that is involved in generating alternative future outcomes. Previous research has highlighted the role of semantic memory in organizing various kinds of future thinking (cf., D’Argembeau & Demblon, 2012; Irish & Piguet, 2013; Klein, 2013; Szpunar et al., 2014), and an important direction for future research will be to examine how semantic memory impacts the generation of alternative future outcomes. Nonetheless, the present research provides clear evidence of a role for episodic memory.

Second, we demonstrated that increasing the number of positive alternative event outcomes was related to larger decreases in the perceived plausibility of the original events. As previously shown in the experiments by Bentz et al. (2009), generating more positive alternative outcomes (long intervention) had no additional effect in decreasing
probability judgments towards a series of negative events, compared with generating fewer alternative outcomes (short intervention), in part due to the more difficult nature of the long intervention which yielded less realistic alternative outcomes. Importantly, here we find that the specificity induction is a useful tool that boosts the perceived ease of simulating these alternative scenarios without decreasing the realistic quality of the imagined events. We were subsequently able to link this boost to larger decreases in the perceived plausibility of the original events when participants generated more alternatives in the specificity condition relative to the control condition. As formerly proposed by a number of researchers (Bentz et al., 2004, 2009; Hirt & Markman, 1995; Hirt et al., 2004), this reduction in pessimistic predictions may be explained via the availability heuristic (Tversky & Kahneman, 1973), whereby certain negative events are more readily available at the forefront of one’s mind, thus increasing likelihood judgments for pessimistic outcomes. Considering more positive outcomes increases the accessibility of and the ease with which one can think about these positive outcomes, relative to the original negative scenarios. In turn, this increased accessibility may reduce the likelihood that individuals will inflate estimates of uncertainty associated with anticipated negative experiences.

Third, we demonstrated for the first time that increasing the number of positive alternative event outcomes is related to larger decreases in the perceived negativity of the original events. These findings may have important implications, given that existing research on affective forecasting has found that individuals are frequently inaccurate when predicting their emotional reactions to negative events (Wilson & Gilbert, 2005). For example, individuals are subject to an impact bias, where they overestimate the intensity and duration of an experience, which may arise from the tendency to focus too much on the event in question and discount the impact of other events or situational factors (Wilson & Gilbert, 2005; Wilson, Wheatley, Meyers, Gilbert, & Axsom, 2000). Gilbert and Wilson (2007, 2009) discuss possible reasons for these errors, including (but not limited to) the fact that the mental “previews” that we generate for anticipated future events: (1) tend to be based on the most available and salient information that comes to mind, which may not be representative of a “typical” experience, (2) may involve a different context than the actually experienced event, (3) generally focus on “essential” features while omitting “incidental” features that are deemed as less important but may actually have a large influence on our subjective emotional experience, and (4) tend to focus on the climax of the emotional experience while disregarding our ability to subsequently adapt to and temper our reactions to such an experience. Accordingly, the process of considering alternative positive outcomes grants access to additional information about ways in which the event may unfold that taps into each of these categories: it increases the availability of positive outcomes, which are most likely comparably plausible to their negative counterparts, it brings to light potential alternative or previously ignored contextual details about the situation, and may call attention to our ability to emotionally and behaviorally adapt to a variety of situations. This suite of additional details may subsequently lead one to realize that the event might not be as predictable or as negative as initially perceived, and thus may encourage a subsequent adjustment of negative expectations.

In the current experiments, we focused on the fluency of alternative event generation and encouraged participants to generate as many alternative event outcomes as possible. However, it is important to note that the concreteness and overall specificity of the alternative scenarios may be another important dimension to examine. It has been reported that individuals with emotional disorders, such as depression (Anderson, Boland, & Garner, 2015; Williams et al., 1996) and anxiety (Brown et al., 2014; McNally, Lasko, Macklin, & Pitman, 1995), tend to show reduced detail and concreteness when retrieving episodic memories and imagining future events. Furthermore, the process of worry in the context of generalized anxiety disorder is thought to involve more verbal and conceptual thought that lacks specific and concrete details, which may elicit less unpleasant physical arousal than visual imagery and episodic simulations (Borkovec et al., 1998; McGowan et al., 2017; Stöber & Borkovec, 2002). It is hypothesized that the more abstract nature of worry allows individuals to disengage and avoid arousing emotional processing towards the anticipated threat, but doing so can have adverse long-term consequences, such as reducing the likelihood that these individuals will imagine their future in a sufficiently concrete fashion to actually cope with the problem (Borkovec et al., 1998; Williams, 2006; Williams et al., 1996). Indeed, recent evidence has shown that increasing the specificity of autobiographical memory and concreteness of mental imagery can be linked to improvements in depressive symptoms (Lang, Blackwell, Harmer, Davison, & Holmes, 2012; Neshat-Doost et al., 2012; Raes, Williams, & Hermans, 2009) and PTSD symptoms (Moradi et al., 2014) with respect to negative and distressing past events, and can also be beneficial for processing worrisome future events that have not yet been experienced (Jing et al., 2016; Skodzik, Leopold, & Ehring, 2017; for review, see Hitchcock, Werner-Seidler, Blackwell, & Dalgleish, 2017). Thus, imagining a smaller number of alternative outcomes in more specific, concrete detail may be just as beneficial in reducing pessimistic future predictions as accessing more alternative outcomes overall, and more research should be conducted to shed light on this point.

It is also important to note that in these experiments we encouraged participants to generate a variety of alternative future scenarios, including changing the actual outcome, the perceived consequences, and other emotional reinterpretations of the original event. However, it is possible that various types of changes may differentially impact one’s perceptions of the event. For example, changing the primary event outcome and the perceived consequences may have a larger influence on the perceived event plausibility and valence than mentally reframing emotional aspects of the original event, or vice versa. Furthermore, certain types of changes may be perceived as more controllable, hence increasing the likelihood of linking mere thought to the actual implementation of action (Epstude et al., 2016). Thus, separately manipulating different types of event changes may be another interesting avenue of future research.

In summary, in these experiments we have advanced our understanding of alternative future event generation by extending the impact of an episodic specificity induction to the simulation of alternative future event outcomes, and also demonstrate that boosting the number of alternative events is related to subsequent changes in the perceived plausibility and negativity of the original events. Further research is needed to characterize the mechanism behind how factors such as the types of alternatives that individuals generate may differentially impact the perceived plausibility and valence of anticipated emotional events, and how imagining alternative outcomes might impact decision-making and the implementation of action. Overall, this line of work may have important implications for understanding how cognitive processes underlying episodic simulation impact the regulation of future-oriented emotion in both healthy and clinical populations.

Acknowledgements

This research was supported by National Institute on Aging grant AG08441 to D.L.S. We thank Laura Hatt, Jenny Sun, and Jyotika Bindra for assistance with scoring participant data.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2017.08.010.

References
