

Trinity University

## Digital Commons @ Trinity

---

Psychology Faculty Research

Psychology Department

---

10-2014

### Interpretive Habit Is Strengthened by Cognitive Bias Modification

Paula T. Hertel

*Trinity University*, [phertel@trinity.edu](mailto:phertel@trinity.edu)

Molly Holmes

*Trinity University*

Amanda Benbow

*Trinity University*

Follow this and additional works at: [https://digitalcommons.trinity.edu/psych\\_faculty](https://digitalcommons.trinity.edu/psych_faculty)



Part of the [Psychology Commons](#)

#### Publication Details

Memory

---

#### Repository Citation

Hertel, P.T., Holmes, M., & Benbow, A. (2014). Interpretive habit is strengthened by cognitive bias modification. *Memory*, 22(7), 737-746. doi: 10.1080/09658211.2013.820326

This Article is brought to you for free and open access by the Psychology Department at Digital Commons @ Trinity. It has been accepted for inclusion in Psychology Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact [jcostanz@trinity.edu](mailto:jcostanz@trinity.edu).

# Interpretive habit is strengthened by cognitive bias modification

Paula T. Hertel, Molly Holmes, and Amanda Benbow

Department of Psychology, Trinity University, San Antonio, TX, USA

(Received 24 March 2013; accepted 25 June 2013)

We investigated the nature of the memory mechanisms underlying cognitive bias modification by applying Jacoby's (1991) process-dissociation procedure to responses during the transfer task. In the two training conditions (negative and benign), students imagined themselves in 100 ambiguous scenarios, most with potentially negative resolutions; the ambiguity was resolved in a consistently negative or benign direction by completing the fragment of a final word. Control participants completed non-ambiguous, non-emotional scenarios. Next all participants responded on a final training block, where half of the scenarios were completed negatively and half benignly. Transfer was assessed by examining choices in the completion of test scenarios when participants were instructed to respond in the same way as they had to a final-block training scenario that was situationally similar. Benign training facilitated correct responding to benign analogues and impaired correct responding to negative analogues. Performance in the negative-training and control conditions was similar. Process-dissociation procedures revealed that this newly established habit and not controlled recollection provided the basis for transfer.

**Keywords:** Cognitive bias modification; Anxiety; Process dissociation; Habit; Transfer.

A negative interpretation bias is a tendency to interpret ambiguous events as emotionally negative or threatening. Consider the example of two students who pass a professor walking across campus. To the student with an emotionally negative bias, the unsmiling nod from the professor means dislike or dismissal; to the student without such a bias, the nod indicates a greeting from someone deep in thought. Anxious individuals are more likely to make negative interpretations of events with ambiguous meaning (Eysenck, Mogg, May, Richards, & Mathews, 1991) and possibly to do so habitually, without apparent awareness of alternative possibilities.

To understand the processes responsible for interpretation biases associated with emotional disorders, biases have been experimentally in-

duced in non-anxious individuals via a procedure called cognitive bias modification (CBM; see Hertel & Mathews, 2011, for review). CBM is achieved through numerous trials that encourage the research participant to resolve ambiguity consistently in one or another direction. One type of CBM paradigm, developed by Mathews and Mackintosh (2000), presents descriptions of scenarios that are emotionally ambiguous until the final word. For example: *For the last few weeks you have noticed that your hearing seems worse than it was. You decide that you should have your ears examined. After a thorough check, the doctor says that your hearing will rapidly:* [word fragment]. When this scenario occurs within a negative training condition of CBM, the upcoming word fragment corresponds to a word that re-

---

Address correspondence to: Professor Paula Hertel, Department of Psychology, Trinity University, 1 Trinity Place, San Antonio, TX 78212, USA. E-mail: [phertel@trinity.edu](mailto:phertel@trinity.edu)

We thank Marqué Moody and Preksha Vankawala for help in data collection.

solves the scenario in a negative manner (e.g., *deteriorate*). When the scenario occurs in a benign condition, its ambiguity is resolved in a non-threatening manner (e.g., with a fragment for *recover*).

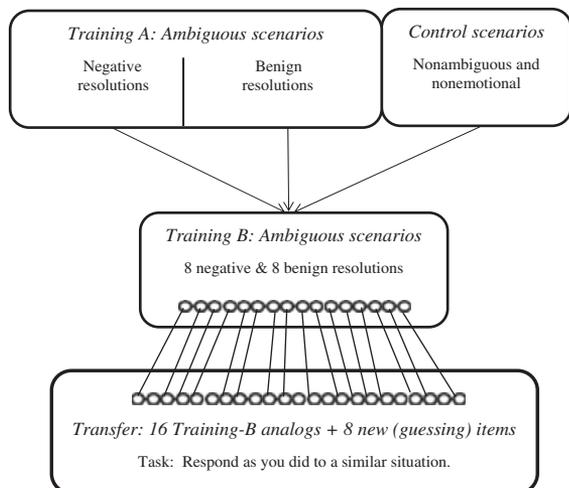
Following a large number of consistently valenced training trials, success in establishing a bias is assessed by examining the effects of training on performance in transfer tasks. Far-transfer tasks reveal effects on reactions to stressful events and are used to argue that cognitive biases indeed establish and maintain anxious states (in the case of negative training) or can be modified to reduce anxiety through benign training (e.g., Wilson, MacLeod, Mathews, & Rutherford, 2006; Woud, Holmes, Postma, Dalgleish, & Mackintosh, 2011). An example of a far transfer stressor is the presentation of video footage from fires and other disasters involving people who ultimately survive. Near-transfer tasks, however, examine whether the bias in question will occur subsequently in the resolution of new ambiguous events that bear some similarity to training (e.g., new ambiguous scenarios). The nature of the near-transfer process is the topic of the present research.

A common assumption is that CBM procedures, like real-life experiences, train habitual ways of responding in new situations—ways that do not depend on the explicit recollection of training events. This assumption seems to have been based on observations that naturally occurring biases in anxious states are not made with deliberate retrospection of prior events but seem to occur without the awareness of alternative possibilities until, for example, they are brought to awareness by therapists. No doubt the habit assumption also derives from the routinisation of the procedure in the training phase of CBM experiments; repetitive practice is crucial to the establishment of new habits. By this reasoning, the transfer task serves as a sort of implicit test of memory or concept learning. It is unlikely, however, that any test of memory or concept learning is process pure (Jacoby, Baker, & Brooks, 1989). The current application of prior experience occurs as some mixture of habitual associative processes and more controlled procedures that focus attention retrospectively. By making assumptions about how habit and controlled recollection concomitantly contribute to performance on a transfer task, it is possible to estimate them separately and then to examine the effects of experimental manipulations on each type (Jacoby, 1991).

We have taken a process-dissociation approach in the current experiment, as well as in two previous experiments that used a different set of instructions (Hertel, Vasquez, Benbow, & Hughes, 2011). Both approaches adjust the typical CBM interpretation paradigm in a way that permits the use of process dissociation procedures. Hertel et al. derived estimates of recollection and habit from performance on transfer scenarios following inclusion and exclusion instructions. On inclusion trials participants were asked to resolve ambiguity by responding in the same way as they had responded to a similar situation encountered at the end of the training phase. On exclusion trials they were asked to remember such a scenario but to respond in the opposite way (with a negative or benign interpretation). The recollection estimates derived from this approach were sensitive to training; compared to both negative training and control conditions, benign training proactively interfered with the recollection of negative resolutions encountered towards the end of training. This was an interesting, and potentially applicable outcome concerning recollection, found in both experiments, but the lack of training effects on estimates of habit ran counter to common assumptions. We therefore sought an experimental approach that would de-emphasise the demand for controlled recollection, believing that effects on estimates of habit would emerge under such conditions. Next we describe our new adjustment to the CBM paradigm and then the current application of process-dissociation equations, modelled on the congruence/ incongruence paradigm used by Jacoby, Debner, and Hay (2001).

In our design the main training phase typified CBM experiments (e.g., Mackintosh, Mathews, Yiend, Ridgeway, & Cook, 2006) but also included 16 additional scenarios at the end of training, half of which were resolved with a negative and half with a benign ending, regardless of the training condition. (We refer to these slightly different training phases as Training A and Training B, and schematise the design in Figure 1.) For the ensuing transfer phase 16 new scenarios were each created to be situationally analogous to one of the Training-B scenarios. The following examples illustrate the match.

(In Training B) *You have agreed to baby-sit while your neighbour goes out for the evening. You have put their daughter to bed but she takes some time to settle down. An hour later you go in to*



**Figure 1.** Schematic representation of the design. Circles represent scenarios; lines represent the situational similarity between Training-B and transfer scenarios.

*check on her and find she is no longer: [fragment for crying or conscious].*

*(In transfer) A neighbour asks you to look after her little girl while she visits a friend in the hospital. The five-year-old cries when her mother leaves but then seems happy to play alone in your garden. After ten minutes you go out and she is: [choice between gone and fine].*

The important feature that distinguished Training-B and transfer scenarios, as is always the case in this CBM/scenario paradigm, was the presence of forced resolutions in the training phase and their absence in the transfer phase. In our transfer phase two alternative endings followed each scenario, and participants were instructed to choose the ending corresponding to the way that they had previously responded to a similar situation. This instruction does encourage a retrospective approach to resolution but does not emphasise recollection to any greater degree than does the typical transfer task used by Mathews and Mackintosh (2000; see also Tran, Hertel, & Joormann, 2011; Woud et al., 2011). It was also the instruction employed on inclusion trials in the experiments by Hertel et al. (2011), with the difference being that exclusion instructions were not used in the current task.

To explain how process dissociation equations are devised to obtain estimates of the role of habit and recollection in this transfer task we use the benign condition as illustration, although the equations would be the same in the negative condition analogously. The first assumption applies to responding when Training A and Training

B are congruent (and benign in our example). When participants respond correctly to transfer scenarios under congruent conditions, they do so to the extent that they recollect the Training-B outcome or, in the absence of its recollection, to the extent that the corresponding habit was trained. Thus:

$$P(\text{correct}_{\text{benign}}) = R + (1 - R) * H \quad (1)$$

where  $R$  represents the contribution of recollection and  $H$  the contribution of habit. On other trials where Training A and Training B are incongruent we are interested in the probability of responding incorrectly with a benign choice. In this case the probability of incorrectly choosing a benign alternative when the similar Training-B scenario was resolved negatively is equal to the probability that the participant responds with a “benign habit” in the absence of recollection. (The incorrect choice would not be made, given recollection). Thus:

$$P(\text{incorrect}_{\text{benign}}) = (1 - R) * H \quad (2)$$

The probabilities of being correct and incorrect in choosing benign alternatives can be estimated in practice by the respective observed proportions of these choices. By subtracting Equation 2 from Equation 1 with those proportions in place, an estimate of recollection is obtained:

$$\hat{R} = \text{prop correct}_{\text{benign}} - \text{prop incorrect}_{\text{benign}} \quad (3)$$

In other words, recollection involves the correct rejection of benign alternatives as well as their correct endorsement. Last, the estimate of habit is calculated by substituting the estimate of recollection into Equation 2.

$$\hat{H} = (\text{prop incorrect}_{\text{benign}}) / (1 - \hat{R}) \quad (4)$$

The habit of responding in a benign fashion, of course, can exist on pre-experimental grounds for someone in any of the training conditions, but is predicted to be stronger in the benign condition due to training.

In summary, process-dissociation procedures operationalise the assumption that habit and recollection jointly determine responding during the transfer phase of interpretation training. This analysis is roughly parallel to the one made by Jacoby et al. (2001) for the purpose of examining memorial differences associated with ageing. Our analysis is done for the purpose of examining the training effects in the CBM paradigm. The training phase can, on both logical and empirical grounds, affect one's ability to recollect individual situations; however, the implicit but compelling assumption of CBM approaches is that training strengthens interpretive habit. In using these new instructions with less emphasis on cognitive control we hypothesised that, compared to a control condition, both negative and benign training would strengthen estimates of those respective habits, although we anticipated that effects might obtain only in the benign condition as had previously been the case (Hertel et al., 2011).

Like Jacoby et al. (2001), but unlike Hertel et al. (2011), we measured guessing in this experiment, by including new test items that did not match up with any Training-B scenarios (see Figure 1). Jacoby et al. (2001) showed that performance on guessing items, assumed to be a relatively pure measure of habit in their experiment, corresponded to their estimates of habit. We therefore expected that habit estimates would not differ significantly from the proportion of guesses. The value of this correspondence for us was that guessing items closely represent transfer items in typical CBM experiments where there is no situational analogy between training and transfer scenarios. Thus, if estimates of habit correspond to performance on guessing trials, the typical presumption regarding the training of habit by CBM procedures is more firmly supported.

## METHOD

### Participants and design

Participants were 72 undergraduate students at Trinity University, pre-screened during their psychology classes to recruit only those with mid-range scores (34 to 47) on the trait form of the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). Students whose scores fell in the first and fourth quartiles were excluded for ethical reasons and to

provide "room" for anxiety scores to reduce following training (see MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002). The students were randomly assigned to the negative, benign, or control training condition (with the constraint of balancing gender and trait scores) and to one of six conditions for counterbalancing materials with (1) their role during the Training-B phase (negative, benign, or missing) and (2) position of the correct alternative during transfer testing (left or right). The final sample of 72 included replacements for 10 students whose data were either incomplete or set aside due to experimenter error (3 students), comprehension problems (3 students), especially large numbers of errors in the training phase (3 students), and previous knowledge of the experiment (1 student). Trait-anxiety scores averaged 40 in each training condition.

### Materials and procedure

*Training A.* In the negative and benign conditions participants read 20 filler scenarios and 80 training scenarios, most of which were adapted from Mackintosh et al. (2006). Each training scenario was identical in the negative and benign conditions and remained ambiguous until the final word. For example: *It is your elderly mother's birthday, and you are going round with some presents. You have the key so you let yourself in, and call out but get no reply. Looking into the kitchen you see she is on the . . .* Participants were asked to read each scenario as if they were the person in the situation and then to anticipate the final word before pressing the spacebar to reveal a fragment to be completed (e.g., *f-oor* for *floor* or *ph-ne* for *phone*). Immediately after typing the word to complete the fragment, participants respond to a "yes or no" question designed to measure comprehension (e.g., *Is your mother talking to someone?*), which would be answered differently, depending on the training condition. The order of presentation was randomised within blocks, each of which contained two fillers and eight training scenarios. An instructional example was discussed at the outset. Fillers were selected from the control scenarios.

In the control condition participants read 100 unambiguous, emotionally neutral scenarios, completed word fragments, and responded to comprehension questions. For example: *You begin reading a book that you recently found around*

*your house. One afternoon you are reading it while sitting in your recliner. You start to feel hungry so you put the book . . .* The spacebar press revealed the fragment (*do-n* for *down*), followed by the question (*Are you planning to take a nap?*). Scenarios were randomly ordered.

*Pilot study for Training-B and test scenarios.* We constructed 24 additional pairs of ambiguous scenarios with the final word missing—one member of each pair to be assigned to the Training-B phase and the other member to the transfer test—and gave them to 35 volunteers to complete with the first word that came to mind. No volunteer completed both members of any pair, and no one participated in the main experiment. Their responses were categorised as benign or negative completions. For counterbalancing purposes in the main experiment the scenarios were organised into six sets of four pairs, and sets were balanced according to the mean proportion of each type of completion. The mean benign completion rate ranged from .56 to .64 for the four scenarios in each set and training/test combination ( $M = .60$ ).

*Training B.* In this phase of the main experiment all participants were presented with 1 scenario from each pair in four of the sets (16 scenarios in total); the scenarios from two sets were accompanied by the fragment for the negative resolution and the scenarios from the other two sets by the benign fragment. The corresponding scenarios final two sets were withheld from the training phase, so that the remaining member of the pair could be used on the transfer test as a control for guessing. Across counterbalancing and training conditions, each scenario was equally likely to be resolved as negative or benign, or to be missing. The 16 Training-B scenarios were presented in a randomised-block order. Each block contained two negative and two benign resolutions. Participants were told to read each scenario carefully because questions would follow.

*Transfer test.* Materials for the test phase consisted of all 24 second-member scenarios from the six sets, 16 of which were situationally similar to the Training-B scenarios and 8 of which were novel. They were presented in randomised blocks of six scenarios, one from each set. We told the participants that each scenario was situationally similar to one of the scenarios they just finished reading in the previous phase: “*For example, if you had just now viewed a scenario*

*about watching a movie, then there would be another scenario about watching a movie in this upcoming set.*” Participants were asked to resolve each scenario in the same way that they resolved the similar scenario in the previous phase and then to press the spacebar to reveal two possible completions (with the position of the correct choice counterbalanced with valence). They chose by striking a key on the number pad (1 for left and 3 for right).

*Other procedural features.* Participants completed mood forms indicating their current state on visual-analogue scales of depressed–happy, tense–relaxed, pessimistic–optimistic, and distressed–not-distressed at three points during the session: prior to the training phase, after the test phase, and after the presentation of short video clips of real, potentially life-threatening situations (Wilson et al., 2006). Participants also completed the state version of the STAI after the third mood form. None of these measures revealed statistically significant differences associated with training and therefore they are not described further.

## RESULTS

### Proportion of correct choices on the transfer test

The first analysis addressed the question of whether training affected correct performance on the transfer test: choice of the negative alternative when the similar Training-B scenario had been negative and the benign alternative when it had been benign. Proportions were submitted to a mixed-design analysis of variance (ANOVA), with a between-participants factor for training condition (negative, benign, or control) and a within-participants factor for valence of completion word in Training B (negative and benign). As the means in Table 1 illustrate, the interaction was significant,  $F(2, 69) = 6.43$ ,  $MSE = .079$ ,  $p = .003$ ,  $\eta_p^2 = .16$ . It was followed by tests of simple main effects of Training-A condition within each type of Training-B completion.

The proportion of correct negative choices differed according to training,  $F(2, 69) = 4.58$ ,  $MSE = .078$ ,  $p = .010$ ,  $\eta_p^2 = .12$ . With the significance level set at .05 (for each set of post-hoc tests in this report), Newman-Keuls tests revealed that benignly trained participants chose correct negative alternatives less frequently than did those in

TABLE 1

Mean (SD) proportion correct during transfer, estimates of recollection and habit, and proportion of guesses

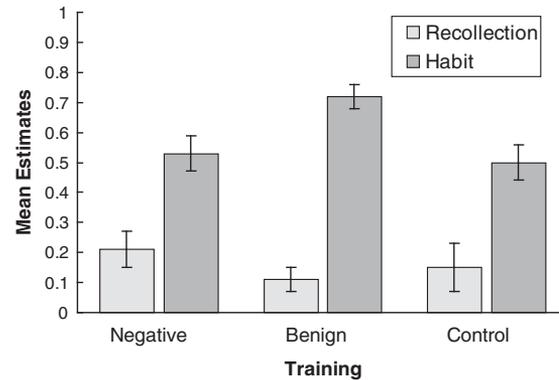
	<i>Negative training A</i>	<i>Benign training A</i>	<i>Controls</i>
Proportion correct			
Benign resolutions	.67 (.18)	.76 (.18)	.58 (.27)
Negative resolutions	.55 (.27)	.35 (.25)	.56 (.31)
Estimates of recollection (for both types of B scenarios)	.21 (.29)	.11 (.18)	.15 (.38)
Estimates of habit			
Incorrect benign/ (1-R)	.53 (.27)	.72 (.22)	.50 (.29)
Incorrect negative/ (1-R)	.47 (.27)	.28 (.22)	.50 (.29)
Proportion guesses			
Benign guesses	.59 (.23)	.77 (.23)	.61 (.28)
Negative guesses	.41 (.23)	.23 (.23)	.39 (.28)

$n = 24$ .

the other two conditions, whose performance did not differ significantly. Thus benign training proactively interfered with memory for negative resolutions, but negative training did not facilitate it. The training conditions also differed in the proportion of correct benign choices,  $F(2, 69) = 4.11$ ,  $MSE = .046$ ,  $p = .020$ ,  $\eta_p^2 = .11$ . According to Newman-Keuls tests, participants in the benign condition chose more correct benign alternatives than did those in the control condition, but not more than those in the negative condition (who did not differ from controls). Thus negative training did not cause interference, but benign training facilitated correct choice.

### Estimates of recollection and habit in transfer

The next set of analyses was performed on estimates of recollection and habit, obtained through process-dissociation equations 3 and 4, to determine the basis of facilitation and interference in the benign condition, reported above. Forced choice between the negative and benign alternatives constrains the analysis to just one type of alternative. Mean estimates for both types, however, are reported in Table 1 to facilitate comprehension of the consequences of the redundancy for each measure. Figure 2 depicts mean estimates of recollection and habit when congruence and incongruence are viewed from the perspective of benign training.



**Figure 2.** Mean estimates of recollection and habit in responding to test scenarios with benign choices. (Error bars represent one standard error.) Estimates of recollection in responding with negative choices are identical; estimates of habit in responding with negative choices are reciprocal of the depicted estimates.

**Recollection.** Estimates of recollection for benign items (calculated as the proportion of incorrect benign responses subtracted from the proportion of correct benign responses) were low and did not differ as a function of training condition,  $p > .49$ . Exactly the same means obtain in calculating recollection of negative resolutions from the perspective of congruence and incongruence with negative training.

**Habit.** Estimates of the habit to interpret in benign ways were calculated as the proportion of incorrect benign responses divided by  $(1 - R)$ , or the proportion of incorrect benign responses adjusted by the absence of recollection. The training effect was significant,  $F(2, 69) = 4.76$ ,  $MSE = .070$ ,  $p = .010$ ,  $\eta_p^2 = .12$ . Estimated habit to make benign interpretations was higher following benign training than in each of the other two conditions according to Newman-Keuls tests. (Reciprocal means result from estimating “negative” habit, as shown in Table 1.)

### Guessing on the transfer test

The proportions of benign resolutions on guessing trials were examined according to training condition. Missing analogues in training, these trials more closely conform to the typical transfer scenarios. The training effect was significant,  $F(2, 69) = 3.43$ ,  $MSE = .062$ ,  $p = .038$ ,  $\eta_p^2 = .09$ . Means are reported in Table 1. Newman-Keuls tests revealed that benign training produced more benign guesses than either of the other two

conditions. To examine whether habit estimates and guessing proportions were similar, we omitted the control data (due to the absence of ambiguity during training). The two measures were non-significantly different,  $p > .182$ , and any tendency towards difference was similar in the two training conditions, as indicated by the non-significant interaction of training condition with measure,  $p > .835$ .

### Training errors

The proportions of errors in resolving the fragments in Training B were infrequent ( $M = .02$ ,  $SD = .05$ ) and did not differ according to training, resolution valence, or their interaction,  $p > .22$ . Errors in resolving fragments during training were also infrequent ( $M = .01$ ,  $SD = .02$ ). However, the proportion of comprehension errors during Training A was relatively high<sup>1</sup> and varied significantly with condition,  $F(2, 69) = 4.99$ ,  $MSE = .005$ ,  $p = .009$ ,  $\eta_p^2 = .13$ . Newman-Keuls tests revealed that the error rate in the negative training condition ( $M = .14$ ,  $SD = .08$ ) was significantly higher than in the control ( $M = .08$ ,  $SD = .06$ ) and benign condition ( $M = .10$ ,  $SD = .06$ ), which did not differ significantly. However, this error rate was not significantly correlated with any of the other performance measures,  $p > .25$ .

## DISCUSSION

This experiment produced three important findings. First, benign training facilitated responding to benign analogues and impaired responding to negative analogues, but negative training failed to have comparable effects. Second, the basis of this transfer effect from benign training emerged in estimates of habit, obtained through process-dissociation procedures (Jacoby et al., 2001), but not in estimates of recollection. Third, the pattern of habit estimates corresponded to the pattern of performance on trials in which guessing was the only means of responding. This third outcome supports the independence claim regarding habit and recollection (see Yonelinas & Jacoby, 2012). Moreover, finding the same pattern in guessing as in habit estimates allays potential concern about

the validity of our approach; guessing trials are comparable to typical transfer trials because there is no situational match.

### Accuracy in near transfer

In two previous experiments conducted with a similar paradigm, Hertel et al. (2011) found impairments in responding to negative analogues following benign training, impairments that they interpreted as proactive interference from benign training in remembering subsequent negative events, but they found only a non-significant trend in facilitating responding to benign analogues. The present results replicated the proactive-interference effect and obtained a significant facilitation effect. However, one subsidiary motivation for the design of the present experiment was the goal of determining whether negative training would exert corresponding interference effects on responses to benign situations; Hertel et al. had not included that combination of conditions. Its inclusion in this design, however, produced no evidence of interference from negative training in remembering benign resolutions as well as no facilitation for negative resolutions. (This lack of facilitation was also found in Experiment 1 from Hertel et al.) Those outcomes resembled performance in the control condition, which did not differ from chance.

More generally, whether effects are achieved by benign or negative training likely depends on pre-experimental tendencies towards one or the other bias; our participants' scores on the STAI-trait form were higher than is typical for student samples, so perhaps their biases were similarly elevated at the start and therefore were more malleable to training in the benign direction. Worth considering in future investigations is the possibility that moderately anxious students are motivated to reject negative-bias training. This possibility is consistent with the higher comprehension-error rates in the negative condition, their slightly higher (but non-significantly so) accuracy rates for benign analogues, and the lack of far-transfer effects on measures of emotion. For the time being, however, this suggestion merely reflects our guess about the absence of evidence for negative interpretation training, evidence often produced through CBM (see Hertel & Mathews, 2011) although there are exceptions that fail to show negative effects with

<sup>1</sup> A review of the comprehension questions revealed that the answers we had selected for several questions were not exclusively correct in the direction assumed.

physically threat (e.g., Mackintosh et al., 2006, Exp. 1).

A possible account of the proportion correct results is the suggestion that participants explicitly respond on the basis of the overall gist of their training (prior to the Training-B scenarios) instead of their memory for specific scenarios. Explicit awareness of the direction of training can account for transfer effects in CBM experiments (Salemink, van den Hout, & Kindt, 2007) and can be seen as clinically instrumental in changing bias. Our experiment did not evaluate the extent of such awareness. If it had been the primary mechanism for responding, however, we would expect strong effects in the negative training condition, because those outcomes were quite salient. The correspondence of performance in the negative and control conditions, even on guessing trials, therefore argues against this account, although it is possible that participants were aware of the nature of their training. It is also important to consider that the effects of overall gist can itself be understood in terms of habit or explicit recollection. As such, an overall gist account begs the question that we sought to answer by using process-dissociation procedures.

### Mechanisms of CBM: Habit versus recollection

The main purpose in conducting this experiment was to investigate the basis of near transfer. Did variations in training affect participants' habitual tendencies, their controlled recollection of analogous scenarios, or both? Hertel et al. (2011) found training effects only on the recollection of Training-B resolutions, whereas we found effects only on estimates of habit. We turn now to a discussion of these differences.

On half the test trials in the experiments by Hertel et al. (2011), habit and recollection were placed in opposition to each other by means of "exclusion" instructions (Jacoby, 1991). Participants were asked to remember the way they had responded to very similar situations in the training phase and then to respond differently (akin to what therapists might instruct their clients to do). Habit would lead them to erroneously respond in the same way as previously, and recollection would lead them to respond differently. In both experiments reported by Hertel et al., recollection of negative outcomes was impaired by benign

training, but the habit to respond in a benign direction was unaffected. In the current experiment we chose a procedure that circumvented the instruction to oppose prior experience and thereby hoped to increase errors in its use. More errors expand the range within which experimental variables such as training valence can exert their effects on habit (with error as numerator).

The elimination of exclusion instructions in the present experiment seemed to have reduced controlled recollection in general, compared to the estimates obtained by Hertel et al. (2011). Moreover, asking participants to respond in the same way as they had previously responded is certainly an instruction more similar to typical CBM-interpretation instructions than is the exclusion instruction. We therefore suggest that the current procedure permits clearer inferences about the processes that are typically invoked by CBM: Benign training strengthens the habit to interpret in non-negative ways and thereby impairs habitual negative biases. And although participants responded with the benign choice correctly more often than incorrectly,<sup>2</sup> the low level of recollection suggests that a carefully attentive focus on the past was not engaged by the instruction to respond in the same way as previously. This interpretation of the difference between the current and previous experiments obviously has not yet been tested. Yet, although there are differences between the samples and procedures of the two experiments, we calculated estimates of recollection from only the inclusion trials in Experiment 2 of the previous report, using the congruence reasoning from the current experiment. The estimate was much higher ( $M = .42$  vs  $.13$  in the same training conditions of the current experiment). This difference ( $p < .001$ ) suggests that inclusion trials are also affected by the degree of control required during exclusion (also see Jacoby, Shimizu, Velanova, & Rhodes, 2005). Moreover, we stress the common thread between the two sets of findings. Benign training makes either habitual or controlled interpretation less negative, depending on the degree to which cognitive control is invoked.

Our procedural analysis rests on the assumption that habit and controlled recollection operate independently in our version of CBM. The preceding suggestion that the detection of habit effects depends on the degree of cognitive control might be compared to the case in which efforts to

<sup>2</sup>Recollection estimates were significantly above zero in both training conditions,  $p < .005$ ; the control condition produced a non-significant trend,  $p = .066$ .

bring a normally automated motor skill under intentional control succeeds in disrupting it. High degrees of cognitive control *seem* at the very least to provide boundary conditions for habitual responding by outshining habit. However, the present instructions and experimental features were designed to achieve conditions for independence and met that test by virtue of similarities of habit estimates to guessing. In future attempts to investigate CBM mechanisms, other assumptions about the relation between habit and control could characterize situations that are designed differently (see Yonelinas & Jacoby, 2012). Our primary goal is to encourage such procedural analyses. Truths arising from the use of different tasks and assumptions might even carry differential implications for clinical treatment.

### Clinical implications

For individuals experiencing certain anxiety disorders (e.g., Generalised Anxiety Disorder or Generalised Social Phobia), treatment addresses biases in the interpretation of ambiguity. Moreover, CBM procedures have recently been developed (e.g., Beard & Amir, 2008). A reasonable speculation about attempts to modify such biases is that they sometimes invite controlled reflective focus on past instances of bias, with the encouragement to now respond differently (exclusion instructions), and they sometimes encourage the practice of new habits in old situations. When cognitive control is minimised (as it is in our current report and probably in typical everyday experience), the trained habit of making benign interpretations determines future responses to ambiguity. And although our transfer task occurred shortly after training, the results tempt us to be confident that more enduring training effects also rest on changes in habit, because controlled recollection should be weakened by longer retention intervals. Alternatively, when recollection is occasionally sought by engaging higher levels of cognitive control, benign training should still prove useful as a means of establishing proactive interference in the controlled recollection of negative interpretations.

One can imagine other clinical situations in which controlled and habitual influences are less independent of each other. In cognitive-behavioural therapy, for example, negative thought records sometimes reveal a primacy of negative interpretation, followed by evaluation of

confirming or disconfirming evidence. When cognitive control operates in this kind of generate/evaluate model, it operates with difficulty. In other words, the training of new habit might trump clinical efforts to encourage re-interpretation or insight. Thus process-oriented laboratory analogues of clinical practice offer the possibility of increasing treatment efficacy or, at least, understanding differential efficacy with greater precision.

### REFERENCES

- Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: Changes in interpretation and social anxiety. *Behaviour Research and Therapy, 46*, 1135–1141.
- Eysenck, M., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*, 144–150.
- Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*, 521–536.
- Hertel, P. T., Vasquez, E., Benbow, A., & Hughes, M. (2011). Recollection is impaired by the modification of interpretation bias. *Journal of Abnormal Psychology, 120*, 902–910.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.
- Jacoby, L. L., Baker, J. G., & Brooks, L. R. (1989). Episodic effects on picture identification: Implications for theories of concept learning and theories of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 275–281.
- Jacoby, L. L., Debner, J. A., & Hay, J. F. (2001). Proactive interferences, accessibility bias, and process dissociations: Valid subjective reports of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 686–700.
- Jacoby, L. L., Shimizu, Y., Velanova, K., & Rhodes, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory and Language, 54*, 493–504.
- Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., & Cook, E. (2006). Induced biases in emotional interpretation influence stress vulnerability and endure despite changes in context. *Behavior Therapy, 37*, 209–222.
- MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*, 107–123.
- Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*, 602–615.

- Salemink, E., van den Hout, M., & Kindt, M. (2007). Trained interpretive bias and anxiety. *Behaviour Research and Therapy, 45*, 329–340.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Tran, T., Hertel, P. T., & Joormann, J. (2011). Cognitive bias modification: Induced interpretive biases affect memory. *Emotion, 11*, 145–152.
- Wilson, E. J., MacLeod, C., Mathews, A., & Rutherford, E. M. (2006). The causal role of interpretive bias in anxiety reactivity. *Journal of Abnormal Psychology, 115*, 103–111.
- Woud, M. L., Holmes, E. A., Postma, P., Dalgleish, T., & Mackintosh, B. (2011). Ameliorating intrusive memories of distressing experiences using computerized reappraisal training. *Emotion, 12*, 778–784.
- Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory & Cognition, 40*, 663–680.