Journal of Cognitive Enhancement
https://doi.org/10.1007/s41465-025-00332-6

ORIGINAL RESEARCH

Applications of Expectancy-Value Theory in Promoting Motivated Behavior in Cognitive Training

Lauren E. Anthony[1] · Aaron Cochrane[2] · Victoria Klaas[1] · Cameron A. Hecht[3] · C. Shawn Green[1]

Received: 11 November 2024 / Accepted: 12 June 2025

Abstract
While the domain of cognitive training is extremely broad, encompassing many different training paradigms and training targets, essentially all paradigms require that participants persist in engaging with a specified training task (or set of tasks) over a long period of time (e.g., tens or even hundreds of hours spaced over weeks or months). As such, in the real world, individuals must sustain motivation to persist in the training of their own volition. Interestingly though, most of the basic science work on cognitive training bypasses the issue of a potential lack of persistence by directly compensating participation (e.g., paying participants). Here we examine the persistence issue through the lens of expectancy-value theory. Participants were provided falsified feedback indicative of different types of performance (e.g., being good at a task from the outset; being able to quickly learn a task) and the impact of those manipulations on persistence behavior was assessed. We found that feedback indicative of improvement was related to longer-term sustained effort, while feedback indicating "good" performance (but no improvement through time) was not. Furthermore, we provide evidence that participants may use local calculations in assessing their current and future improvement, which presents challenges for providing feedback that is both accurate and motivating in the context of cognitive training paradigms.

Keywords Cognitive training · Expectancy-value theory · Motivation · Persistence

Introduction

The past several decades have seen a significant increase in research focused on the enhancement of cognitive abilities via dedicated behavioral cognitive training (Deveau et al., 2015; Green & Bavelier, 2015; Guye et al., 2021; Katz et al., 2021; Schubert et al., 2014). While the entirety of the literature encompasses studies employing an extremely wide range of training approaches and a similarly broad range of targeted cognitive domains (Green et al., 2019), the available empirical data provide reason to be hopeful that some forms of behavioral training may be viable avenues to consistently augment at least some cognitive functions.

The area of greatest debate in the field (e.g., Simons et al., 2016), and thus perhaps not surprisingly, the most research, has focused on the question of whether various forms of behavioral training produce generalizable increases in cognitive function and, if so, how best to design training for generalization (Deveau & Seitz, 2014; Green & Bavelier, 2008; Pasqualotto et al., 2023). One key theme that has emerged from this work is that very short amounts of training, for instance, just a few hours, are unlikely to produce changes in broad cognitive functions. Instead, much longer training durations spread out over weeks or months are likely required to produce enhancements at the level of cognitive constructs (Bediou et al., 2023; Schmiedek et al., 2010). In the real world, participants will thus need to stay self-motivated to continue using these paradigms long-term. Yet, in most long-term intervention studies designed to examine the impact of various forms of behavioral training, the participants have been paid (or otherwise compensated) to take part in the studies. It is frequently even the case that, in order to reduce attrition, participants are both paid for their regular participation and also offered a bonus for completing the

✉ Lauren E. Anthony
leanthony2@wisc.edu

1 Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

2 Department of Psychology, Brown University, Providence, RI, USA

3 Department of Psychology, University of Rochester, Rochester, NY, USA

Published online: 20 June 2025

full study. Such a set of compensation strategies is eminently sensible given a primary focus on questions related to generalization and the desire to (A) ensure that a random sample of participants is recruited (i.e., to not just recruit participants who might be internally motivated to engage long-term with these forms of training), and (B) to ensure that most participants complete the training and thus have their data included in final analyses. However, this is also a weakness, since ensuring that participants have the internal motivation to persist with long-term training is likely to be a significant issue, particularly in cases where the training is dull or repetitive. Although many forms of behavioral training for cognitive enhancement have attempted to insert "game-like" elements to reduce the dullness/repetitiveness, not only does this not always work, there is evidence to suggest that some instantiations may be problematic and perhaps even counterproductive (Vermeir et al., 2020). For instance, inserting gamification features such as real-time scoring and scaffolding both led to negative impacts on training performance (Katz et al., 2014). This suggests the need for more basic science foundations to be built with regard to keeping participants engaged for the long-term in cognitive training.

Although the problem of keeping participants engaged over the long-term has not been a primary focus of cognitive training research thus far, there is a great deal of work in other domains examining the factors that impact individuals' motivation to persist with an activity. For example, Eccles' expectancy-value theory of motivation is one of the major theories that has been developed to predict and understand individuals' task engagement decisions (Eccles & Wigfield, 2020; Eccles et al., 1983). According to this theory, individuals are more likely to engage in a task when they (1) value the task (i.e., when they believe that it is useful or important or when they enjoy the task) and (2) believe that they can succeed at the task. A large body of correlational and longitudinal research is consistent with this theory (e.g., Harackiewicz et al., 2008; Jacobs et al., 2002; Wang, 2012). For instance, research has shown that students who perceived high levels of expectancy and value for math and science were more likely to continue enrolling in these types of courses (Simpkins et al., 2006). Likewise, students who believed that they could be successful at math were more likely to continue enrolling in math courses.

In addition, a growing body of research suggests that the effects of expectancy for success and value are not merely additive, but multiplicative (Guo et al., 2016; Nagengast et al., 2011). Specifically, this research suggests that individuals must *both* perceive value in the task and expect to be successful to be optimally motivated. For example, if an individual values a task but does not believe that they can be successful, they may see little point in engaging in the task. Conversely, if an individual believes that they can succeed in a task but sees no value in succeeding, they may have little motivation to engage in the task.

In the context of cognitive training tasks, one possible way to influence individuals' motivation to persist in training is to selectively spotlight, alter, obscure, or even falsify the presentation of the feedback they receive about their performance. For instance, feedback can emphasize individuals' effectiveness at the outset of the task (e.g., after the first few trials) and/or can emphasize participants' improvement on the task. Here there are multiple possibilities regarding exactly how these two manipulations (emphasize current skill or emphasize improvement) may influence individuals' motivation to persist with a cognitive training task. On the one hand, the two manipulations might be equally effective at increasing motivated effort, given that they should both increase individuals' expectations for success. Giving participants positive feedback about their performance at the outset of a task should increase their belief that they can be successful at the task going forward (or, indeed, indicate that they are already successful). Similarly, giving participants feedback about their improvement throughout the task should also increase their expectancies for success by indicating that they are likely to become very good at the task in time.

On the other hand, giving participants positive feedback at the outset of a task—if not complemented by information about improvement over time—may actually *undermine* motivation to persist at the task. The reason for this is that if an individual feels that they have mastered a task right away, they may perceive that there is little value in persisting. That is, subsequent trials may fail to yield any additional perceived value (e.g., cognitive improvements) if an individual has already reached the peak of their performance. If this is the case, then even with heightened expectations for success, the individual might be unlikely to persist with the task because doing so would lack value.

The Present Research

Here we tested these competing hypotheses by manipulating participants' beliefs about their ability and improvement on cognitive training tasks via false feedback. To do so, we first examined a critical, yet reasonably untested, question—whether individuals have accurate beliefs about their abilities in the context of a standard cognitive training task. If individuals have reasonably accurate beliefs about their abilities, it would be difficult to manipulate those beliefs via falsified feedback (because they would notice a mismatch between how they believed they performed and the feedback). Thus, in an initial experiment, we examined the accuracy of individuals' beliefs about their current and future ability on a standard task utilized in many forms of cognitive training—the N-back task. As our findings reveal, individuals in fact have a reasonably poor estimate of their own abilities, at least in the absence of explicit feedback, thus making deliberate manipulations of feedback a viable avenue.

We then manipulated beliefs about current and future ability by providing individuals with false feedback about their performance in order to examine how these manipulations impacted the decision to persist with the training. Specifically, we contrasted the impact of two main types of manipulations in terms of their ability to induce motivation to persist: (a) feedback suggesting that the individual was already good at the task (but was not improving) or (b) feedback suggesting that the individual was initially poor at the task but was rapidly improving. To preface the key result, we found the latter manipulation—feedback indicating improvement—was far more powerful in terms of producing motivated effort to continue on the N-back task than feedback indicative of initially good (but not improving) performance. This result has clear implications for the design of feedback in cognitive training paradigms utilized in applied real-world settings, which we consider more fully in the discussion.

Study 1

Before considering the impact of purposefully manipulating expectancy-based beliefs about cognitive training task performance on motivated effort, it is first necessary to examine the extent to which individuals have an accurate internal estimate of their abilities on that task. If individuals can accurately assess their own ability, they may be less likely to accept falsified feedback about their performance. If, however, individuals' internal understanding of their performance/ability to improve is poor, they may be more accepting of falsified feedback.

In Study 1, participants underwent four sessions of experience with a spatial N-back task. Prior to starting each session, they were given a survey asking them to indicate their beliefs about their current and future performance levels. We then examined the extent to which the participants' stated expectations regarding their performance matched their true levels of performance. Our expectation was that there would be a poor initial match between their internal beliefs about their abilities and their true abilities, but then, when given accurate feedback about their performance, their estimates would come to match that feedback.

Materials and Methods

Participants

Twenty-two participants were recruited either from an Introductory Psychology course or via on-campus posters (M_{age} = 19.79 years +/− 2.19 years; 13 females). As race/ethnicity data were not collected for any of the studies presented in this report, here we provide the composition of the University of Wisconsin-Madison at the time the studies were conducted (2017–2018): 71–73% White, 6% Asian, 5% Hispanic, 2% Black, < 1% Indian, < 1% Native Hawaiian, 3% multi-racial (non-Hispanic), and ~ 1% unknown. Those recruited from the Introductory Psychology course participated in the study in exchange for course credit, while those recruited via on-campus posters were paid $10/h for their participation. All participants gave informed consent to participate in this research. There were no statistically significant differences between participants recruited via different means on either N-back performance or expected performance; however, we note that the number of paid participants was quite small and thus this was an underpowered analysis.

Human and Animal Rights

All research activities were approved by the University of Wisconsin–Madison Institutional Review Board.

Apparatus

All tasks were presented on 22-inch LCD monitors using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007) in MATLAB, running on Windows computers.

Task and Procedure

Overview The full experiment consisted of a total of 4 sessions, each held on a separate day, with all sessions being completed over a maximum of 10 total days. On the first day of the study, after being given a broad overview of the study and providing informed consent to proceed in the study, participants were given a verbal and written description of the upcoming N-back task, which also included a short paragraph suggesting the possible value of the N-back training task. The paragraph defined working memory and informed participants that training working memory can improve intelligence and is associated with improved life outcomes (for full value-added statement, see supplementary materials). Participants then completed a survey in which they were asked about their expectations about the upcoming N-back task. Most relevantly, they were probed as to their expectations of their upcoming performance (e.g., "How good do you think you will be on the N-back when you start today?"; see supplementary materials for a full list of surveys and questions). This was then followed by 30 minutes of training on the N-back task. The three subsequent sessions followed the same pattern but without the initial verbal description of the study/task or the individual difference measures (i.e., sessions 2–4 started with participants completing the assessment of their expected performance on the N-back in the upcoming session followed by 30 minutes of the N-back task). On the final day of training, participants also completed an additional survey following the N-back task asking them to predict how they would have performed on the N-back task if they were to complete an additional training session.

N-Back Training Task The N-back task that participants completed during each session was a standard spatial N-back task (4-back) (Jaeggi et al., 2008). On each block of this task, participants saw a series of 30 squares presented one after another in one of eight possible locations on the computer screen. Starting with the 5th square that was presented, participants were asked to indicate whether the square they were currently viewing was in the same spatial location as the square they saw 4 trials previously (e.g., for the 5th square, was it in the same position as the 1st square that was presented?). After every block (i.e., the full set of 30 squares presented), participants received feedback as to their overall percent correct for their responses on that block. Each session lasted for 30 minutes with participants completing as many blocks as fit in that time window (mean number of blocks completed per session = 25.1 +/− 5.8).

Results

Exclusions

A total of 20/22 participants completed between 83 and 126 blocks in total across the 4 sessions with the remaining 2 participants completing only 44 and 41 blocks respectively. These 2 participants did not appear to be following overall task instructions and were removed from further analyses.

Analyses

We first examined the extent to which participants' expectations about their ability (i.e., predicted percent correct) were related to their actual ability at the start of the task. As expected, there was not a reliable correlation between participants' initial expectations of performance on the first day and subsequent task accuracy on the first day ($r = 0.03$, $CI_{95\_bootstrap} = [-0.68, 0.53]$; see Fig. 1).

We then asked whether participants' understanding of their ability improved through time as they received feedback about their performance. Again, as expected, given (accurate) feedback about their performance, participants' ability to predict their upcoming performance gradually improved through time (Fig. 2).

The monotonic effect of day was reliably predictive of median absolute prediction error (in a Bayesian mixed-effects model including by-participant monotonic effects of day on median absolute error), indicating an average per-day decrease in prediction error of −2.01 ($CI_{95} = [-3.75, -0.32]$). This was also confirmed by testing the product-moment correlation between last-day prediction and last-day performance ($r = 0.79$ $CI_{95\_bootstrap} = [0.51, 0.92]$.

In contrast, the effect of day was not reliably predictive of individuals' evaluations of their most recently completed block, though this may be due to the generally low errors across all days (estimated per day change −0.30, $CI_{95} = [-1.12, 0.56]$; see Fig. 3).

Discussion

Consistent with our expectations, participants' expected performance on the N-back task prior to experiencing the task and receiving feedback about their performance was

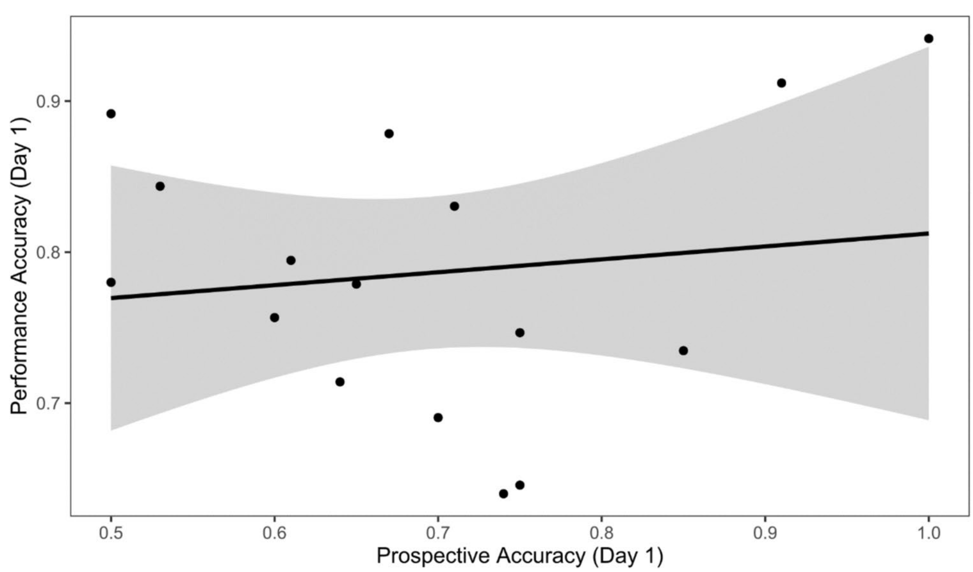

Fig. 1 Day-1 prospective accuracy expectation and actual accuracy, with linear best-fit line and 95% CI. As is clear, there is no evidence of a positive correlation between participants' prospective expectations of performance and actual accuracy on the N-back task. Instead, it appears that participants are not able to accurately estimate how well they will perform on the N-back task

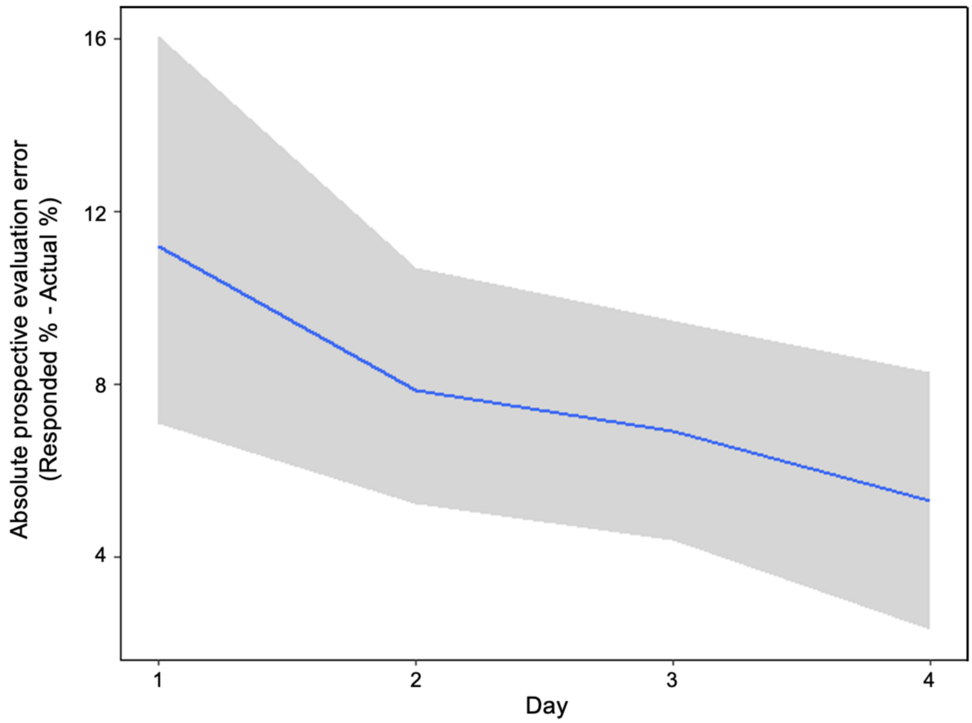

Fig. 2 Prospective evaluation error across days. Monotonic regression estimates and 95% confidence intervals indicate a substantial decrease in participants' median absolute error when asked to predict their performance on an upcoming block. In other words, estimates of prospective performance became more accurate across experimental sessions

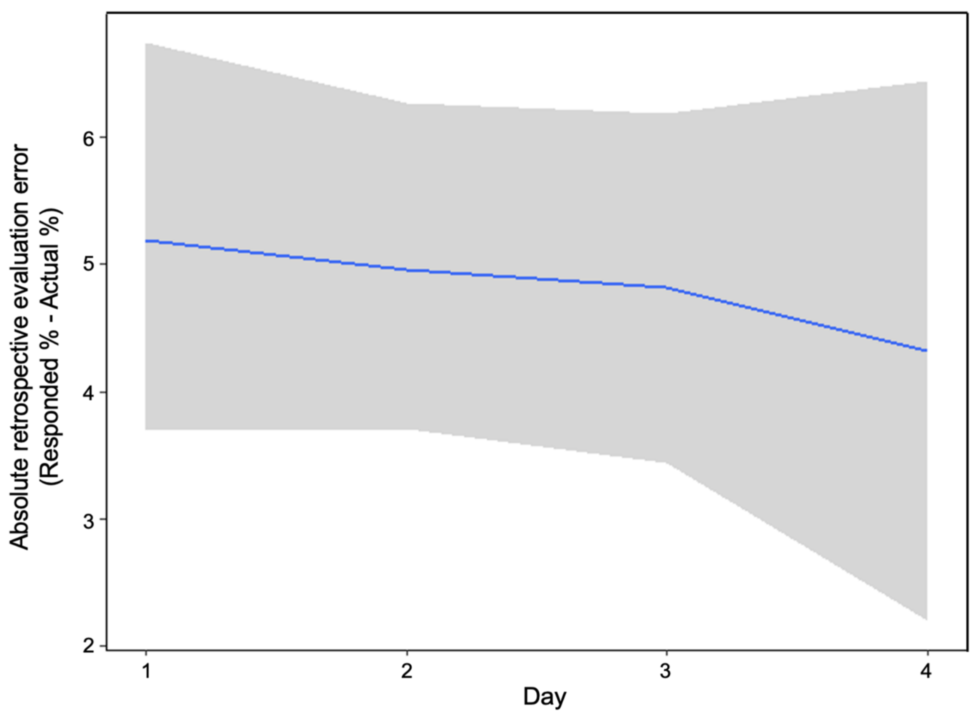

Fig. 3 Retrospective evaluation error across days. Monotonic regression estimates and 95% confidence intervals indicate stability in participants' median error when asked to estimate their performance on a preceding block

essentially unrelated to their actual level of performance. Through experience and accurate feedback, participants' expectations of their upcoming performance became progressively more accurate. The fact that participants had a reasonably poor initial understanding of their ability thus opens the door for providing inaccurate feedback as a deliberate manipulation to increase and/or decrease the motivation to persist with the task. This will be the focus of Studies 2 and 3 below.

Study 2

Although in Study 1 we saw that participants had a poor understanding of their ability in the N-back task, their level of understanding was not zero. Thus, rather than launching directly into a study utilizing falsified feedback to alter persistence motivation in the context of an N-back task, we instead chose to examine the general to-be-utilized approach to altering motivated behavior in a task where there was no apparent way for participants to gauge their own performance level outside of direct feedback. This would allow us to examine whether the main manipulations of interest did in fact produce the changes in persistence behavior that were anticipated. As noted above, we sought to examine whether providing false feedback indicating that individuals (a) were immediately good at the task and/or (b) were rapidly improving at the task increased persistence and whether one or the other of these beliefs more strongly impacted persistence behavior.

In brief, in Study 2, participants took part in a "categorization task" where they were asked on each trial to indicate which of four possible categories a particular abstract shape belonged to. In actuality, there was no category structure to the shapes (i.e., stimuli were drawn from a uniform distribution and thus even an unsupervised learning system could not "do" the task accurately) and all feedback was predetermined. Specifically, participants were randomly assigned to receive block-wise (every 40 trials) feedback indicating that they were either: poor at the task from the beginning and did not improve further (i.e., low initial skill, low improvement), excellent at the task from the beginning, but then did not improve further (i.e., high initial skill, low improvement), or poor at the task from the beginning, but then improved at the task rapidly (i.e., low initial skill, high improvement). After each block, participants were asked whether they wanted to quit the current task and switch to something new. The critical question was thus whether the type of performance feedback they were provided impacted when/if participants chose to quit the task.

As discussed above, we tested two competing possibilities. The first possibility was that participants would be equally likely to persist when told that they were good at the task from the beginning but without further improvement (high initial skill, low improvement) and when they were told that they were initially poor at the task but improved rapidly (low initial skill, high improvement), because each of these feedback manipulations would increase participants' expectancies for success. The second possibility was that participants would be more likely to persist when told that they were improving than when told that they were initially good at the task but not improving because the latter feedback manipulation would reduce the value of persisting with the task.

Materials and Methods

Participants

One hundred one participants enrolled in an Introductory Psychology course participated in the study in exchange for course credit (M_{age} = 19.22 +/− 1.42; 56 female). All participants gave informed consent to participate in this research. No participants who participated in Study 1 were recruited to participate in Study 2. Participants underwent debriefing at the end of the study, wherein the full design, including the use of and justification for deception, was explained, and participants were given opportunities to ask questions.

Human and Animal Rights

All research activities were approved by the University of Wisconsin–Madison Institutional Review Board.

Apparatus

All tasks were presented on 22-inch LCD monitors using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007) in MATLAB, running on Windows computers.

Task and Procedure

The basic stimuli were similar to those utilized in Kattner et al. (2017a) and Kattner et al. (2016). On each trial, a simple line drawing was presented on a gray background. These stimuli were referred to as "feathers" to the participants. The feathers consisted of a central vertical line with oriented lines branching off symmetrically to the left and the right. The length of the central vertical line, the branches, the orientation of the branches, and the number of branches were drawn from uniform random distributions on each trial. Participants were told that there were 4 "types" of feathers and that their job was to select the correct category for each feather they viewed. Participants indicated their choice by clicking on one of four differently colored horizontal lines. There were no normatively/predefined "correct" responses in this case, as the stimuli were simply randomly drawn from the full stimulus space. However, participants were unaware of this fact. Participants received "feedback" in the form of a graph at the conclusion of each block of 40 trials (Fig. 4). This feedback was pre-defined based upon their group assignment (see below) and thus did not in any way correspond to their actual responses. After the completion of the third block and each subsequent block, participants were asked the following, "Would you like to continue classifying

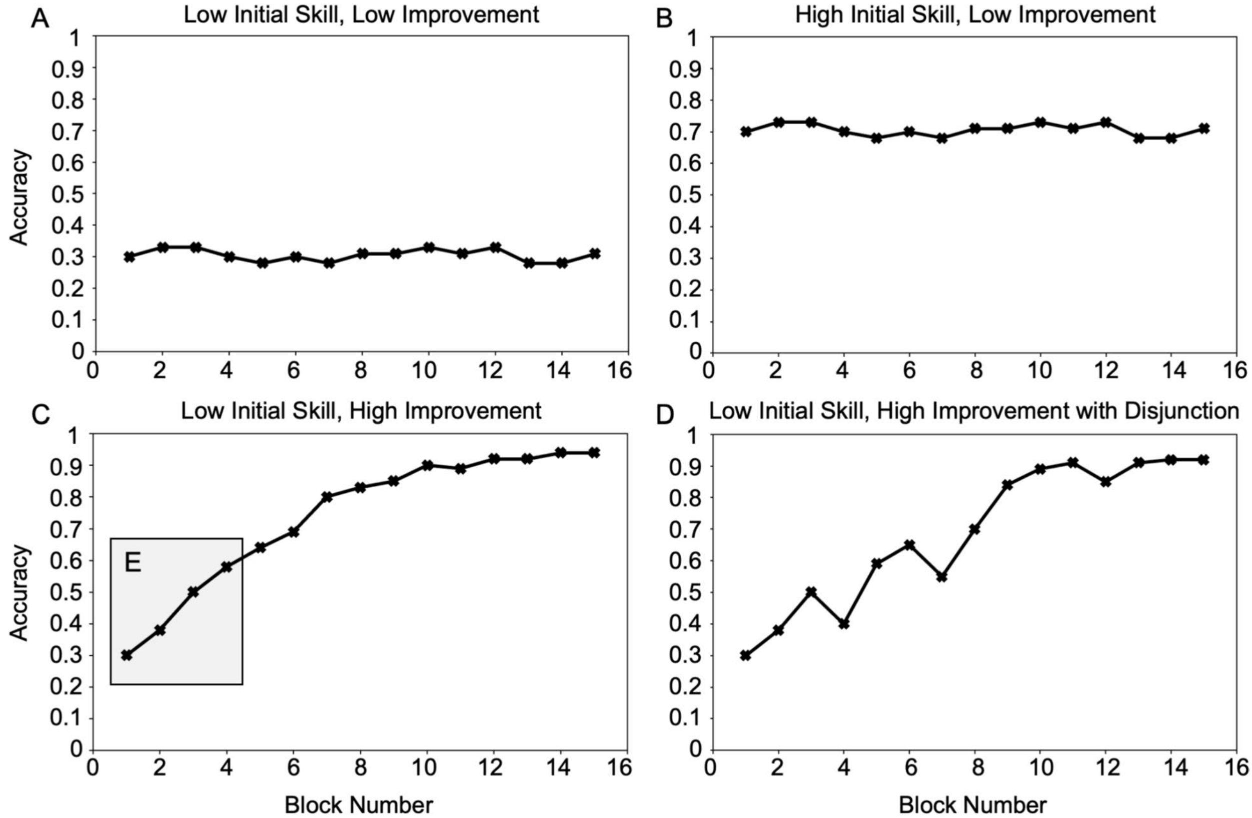

Fig. 4 Participant feedback graphs. **A** Feedback graph shown to participants in the low initial skill, low improvement condition, **B** high initial skill, low improvement condition, **C** low initial skill, high improvement condition, and **D** low initial skill, high improvement condition with disjunction condition. **E** Participants were presented with a feedback graph after each block. For example, after block 4, participants in the low initial skill, high improvement condition were shown graph **C** with only the first four points plotted

the first set of feathers or switch to the second set of feathers?" The key dependent measure was thus the block after which they decided to quit the initial feather task and switch to the second feather task. Following the completion of 14 total blocks (feather task 1 + feather task 2), participants were informed that the study was completed.

Feedback Conditions Participants were randomly assigned to one of four feedback conditions. All aspects of the procedure were identical across conditions, with the exception of the feedback graph that was provided to participants after each block of trials. Three of the conditions corresponded to conditions that could be used to evaluate clear hypotheses from expectancy-value theory. These conditions, in turn, were labeled: "low initial skill, low improvement," "high initial skill, low improvement," and "low initial skill, high improvement" (see Fig. 4). We examined whether the "high initial skill, low improvement" and the "low initial skill, high improvement" conditions increased persistence as compared to the "low initial skill, low improvement" condition. Note that it was not possible to have a "high initial skill, high improvement" group, as the dependent measure provided to participants as feedback was accuracy and thus was bounded at 100%. This issue was addressed in Study 3.

We further note that there was a fourth condition that participants were assigned to which was not part of the initial planning for the study. When the experiment was initially conceived and implemented, in an attempt to make the "low initial skill, high improvement" feedback plot look "plausible", several disjunctions were put into the graph (i.e., where performance on the current block was indicated to be worse than in the previous block). The first of these disjunctions occurred for the feedback that was provided after block 4. We thus refer to this as the "low initial skill, high improvement (with disjunction)" condition. This is treated as a separate condition due to the fact that early in data collection, it became obvious that the presence of the disjunction on block 4 was dramatically impacting persistence behavior (specifically, a substantial number of individuals quit directly after receiving the feedback in block 4). As this was a potentially interesting finding, we continued collecting data on this version while also utilizing a different, monotonically increasing version of the feedback graph for what became the final "low initial skill, high improvement" condition. Consistent with best reporting practices in the behavioral sciences, it is important to note that the hypotheses that emerged around the disjunction condition were a consequence of early data collection, rather than being pre-specified. This issue was rectified in Study 3.

Results

Exclusions

One of the 101 participants indicated that they did not believe the feedback was accurate (and in fact, went so far as to deliberately, in one block, choose randomly and in another block choose the opposite of what they would have normally chosen to show that the feedback was not linked to their choices). This single participant was thus removed from further analysis.

Analyses

Our main questions here were with respect to whether participants, on average, chose to quit the task earlier/later depending on the feedback condition to which they were assigned. We thus used what is commonly referred to as event history analysis in the social sciences (or survival analysis in the biological/medical sciences). Here an "event" is defined as the point at which a participant decides to switch tasks.

First, although most of our comparisons were planned and were with respect to pairwise effects (e.g., one condition versus another), in order to be conservative, we chose to first conduct an omnibus test that included all four conditions, because the comparisons of interest constituted non-orthogonal contrasts (i.e., we were unable to conduct all of the desired pairwise comparisons using orthogonal contrasts) (Ruxton & Beauchamp, 2008). The results of the omnibus test were significant ($X^2(3) = 15.43$, $p = 0.002$); this test was conducted with all four conditions. Given the significance of the omnibus test, we felt confident in proceeding with our planned comparisons.

Next, we compared our three primary conditions of interest: "low initial skill, low improvement," "low initial skill, high improvement," and "high initial skill, low improvement." Those participants who never decided to switch tasks were labeled as "censored" (i.e., data were "right" censored). The log-rank test indicated that these groups were significantly different ($p = 0.001$; The log-rank test tests the null hypothesis that there is no difference between groups in the probability of an event at any time point; Bland & Altman, 2004). In other words, the difference in survival probabilities between "low initial skill, low improvement," "low initial skill, high improvement," and "high initial skill, low improvement" conditions was significant. Kaplan–Meier curves were estimated for these three groups and are plotted in Fig. 5.

We next pairwise compared specific conditions to investigate the effects of feedback about initial performance (i.e.,

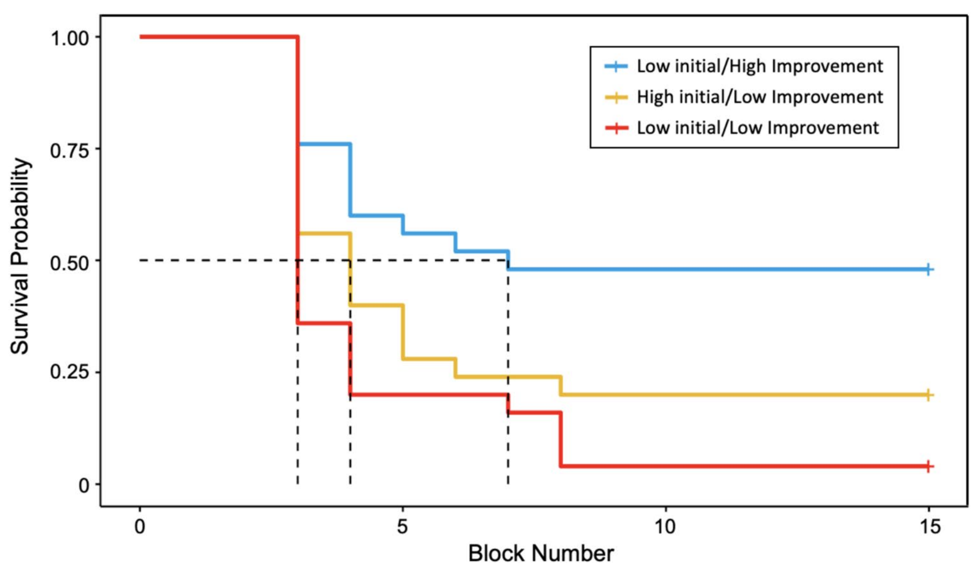

Fig. 5 Kaplan–Meier curves estimated for each of the three main experimental conditions. The significance of the log-rank tests suggests that the survival probabilities of these three conditions are different. Specifically, participants quit most quickly when they received feedback that they were performing poorly and not improving (low initial skill, low improvement) as compared to when they received feedback that they were performing well, but not improving (high initial skill, low improvement) or to when they received feedback that they were initially performing poorly but were rapidly improving (low initial skill, high improvement)

high or low) and the effects of feedback about performance rate (i.e., high or low) separately in a Cox proportional hazards regression analysis. To make these comparisons, our condition variable was dummy coded to create three non-orthogonal contrasts; the "low initial skill, low improvement" condition was specified as the reference group in this analysis. First, to assess the effect of feedback about initial performance, we examined the parameter estimate associated with the "high initial skill, low improvement" and the "low initial skill, low improvement" group comparison. Here, although the data were in the predicted direction, wherein the high initial group had a decreased risk of switching tasks relative to the low initial group, this effect did not reach significance in a two-tailed test (Hazard Ratio: 0.59, CI_{95} [0.32,1.07], $p = 0.082$) (Note: a hazard ratio less than 1 indicates decreased risk, a hazard ratio equal to 1 indicates a lack of association between condition/treatment assignment and risk, and a hazard ratio greater than 1 indicates increased risk). Additionally, the log-rank test indicated that these groups were not significantly different ($p = 0.12$).

Next, we investigated whether the feedback that was provided to participants about rate of performance affected the probability of dropping out (i.e., choosing to switch to a new task). For this comparison, our condition variable was recoded to create three non-orthogonal contrasts with the "low initial skill, high improvement" condition specified as the reference group. To assess the effect of feedback about rate of performance improvement we examined the parameter estimate associated with the "low initial skill, high improvement" and the "low initial skill, low improvement" group comparison. The direction of the effect was such that the low improvement group had an increased risk of switching tasks relative to the high improvement group. Specifically, participants in the low improvement condition were 3.71 times more likely to choose to switch tasks than those in the high improvement condition (Hazard Ratio: 3.71, CI_{95} [1.88,7.34], $p < 0.001$). Likewise, the results of the log-rank test indicated that these two groups were significantly different ($p < 0.001$).

The last analysis conducted was an unplanned comparison between the "low initial skill, high improvement" and "low initial skill, high improvement (with disjunction)" conditions (*note: we consider this an "unplanned" comparison because the disjunction group itself was not part of our a priori plans). As discussed above, although our original intention with the feedback graph that was shown to the "with disjunction" group was to provide "realistic" feedback where performance did not neatly and monotonically increase through time (i.e., to guard against participants distrusting the accuracy of the feedback), we noticed that a significant number of participants chose to switch tasks upon reaching the first point of disjunction (i.e., the first point where they were told their accuracy rate had decreased). Thus, we were interested in comparing each of these conditions to assess whether the introduction of these disjunctions resulted in significantly increased risk of task switching. We used the same three non-orthogonal contrasts as in the analysis above with the "low initial skill, high improvement" condition specified as the reference group. Descriptively, the "low initial skill, high improvement (with disjunction)" group was at an increased risk of switching tasks relative to the "low initial skill, high improvement (without disjunction)" group, but this difference was not significant (Hazard Ratio: 1.89, CI_{95} [0.94, 3.79], $p = 0.074$). Additionally, the log-rank test indicated that these groups were not significantly different ($p = 0.095$). Kaplan–Meier curves were estimated for this comparison and are plotted in Fig. 6. Although these effects did not reach significance, given that

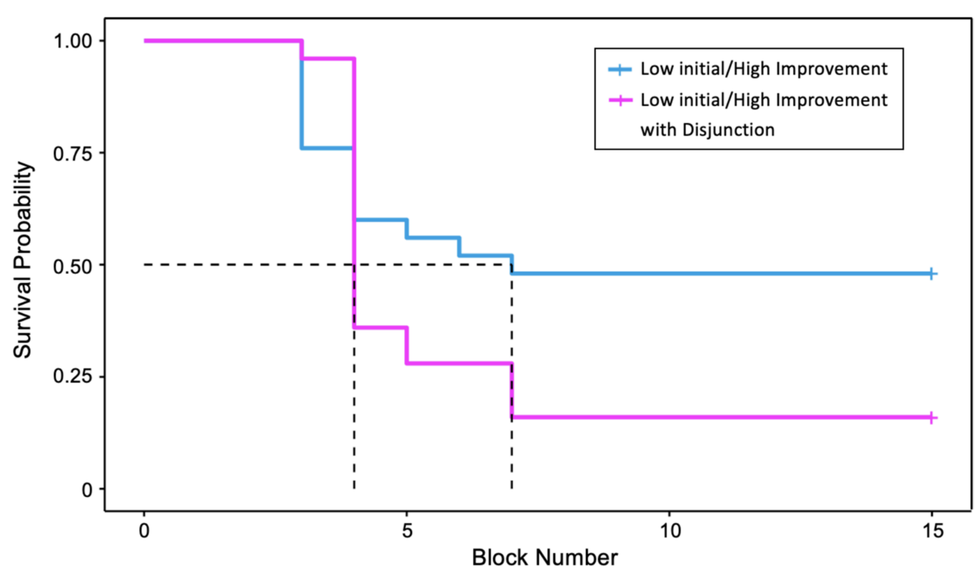

Fig. 6 Comparison between the low initial-high improvement and low initial-high improvement with disjunction groups. Kaplan–Meier curves estimated for each of these conditions (log-rank test $p = 0.095$). Descriptively, survival probabilities differ between these conditions, but the difference did not reach significance

they were unplanned in Study 2, and given the strong direction of the effect, we chose to examine the difference between "low initial skill, high improvement (with disjunction)" and "low initial skill, high improvement (without disjunction)" in a planned comparison in Study 3.

Discussion

Overall, we found strong support for the hypothesis that providing different feedback regarding the rate of improvement would alter persistence. Specifically, participants given feedback that they were rapidly improving at the task persisted in the task to a far greater degree than participants who were given feedback that they were not improving at the task. Importantly, we found essentially no evidence that feedback related to overall performance levels impacted persistence. Participants who were provided feedback that they were performing at a very high level (but not improving) showed no greater persistence in the task than did participants who were provided feedback that they were performing at a very low level (but again, not improving). This finding was consistent with the hypothesis that participants may not persist at a task, even when they expect to succeed at it, if they do not perceive that they are improving (particularly when they do not find the task to be intrinsically valuable or enjoyable). In such cases, individuals may perceive that there is little value in continuing because they have already reaped all of the benefits they are likely to receive from the task.

Finally, although it did not reach the level of significance, we found evidence that was at least suggestive of the idea that participants were utilizing a local calculation of improvement in making their decision to stay or switch tasks. Participants given feedback that they were globally improving, but where their most recent block performance was lower than the one previous, tended to quit earlier than participants who were given feedback that they were monotonically improving.

Study 3

In Study 3, we built on the results of Study 2 by using the same basic feedback manipulations in the context of a typical working memory task. Furthermore, given the unexpected finding that participants may be utilizing very local information in assessing their degree of improvement, we chose to a priori include a condition that had similar local disjunctions in the feedback graph as in Study 2.

Materials and Methods

Participants

One hundred sixteen participants enrolled in an Introductory Psychology course participated in the study in exchange for course credit (M_{age} = 19.40 +/− 1.12; 70 female). All participants gave informed consent to participate in this research. No participants who participated in Studies 1 or 2 were recruited to participate in Study 3. Participants underwent debriefing at the end of the study wherein the full design, including the use of and justification for deception, was explained, and participants were given opportunities to ask questions.

Human and Animal Rights

All research activities were approved by the University of Wisconsin-Madison Institutional Review Board.

Apparatus

All tasks were presented on 22-inch LCD monitors using Psychtoolbox (Brainard, 1997; Kleiner et al., 2007) in MATLAB, running on Windows computers.

Task and Procedure

The basic study design in Study 3 was identical to that in Study 2 with the exception of the initial training task and the to-be-switched to task. First, rather than performing the feathers categorization task, participants initially performed the same spatial N-back task as in Study 1. When they were given the option to change tasks starting after block 3, this was phrased as: "Would you like to continue with the N-back task or switch to a different cognitive task?" If they decided to switch tasks, they changed to a simple reaction time task (performance on this latter task was not analyzed or considered further).

Feedback Conditions Participants were randomly assigned to one of five feedback conditions. Given the results of Study 2, we included a "low initial skill, high improvement (with disjunction)" condition as part of our a priori condition set (with the expectation that this condition would produce less persistence than the "low initial skill, high improvement" condition). Furthermore, in Study 2, feedback was provided in terms of accuracy, which was necessarily capped at 100%. This eliminated our ability to utilize a "high initial skill, high improvement condition." In Study 3, in order to rectify this issue, performance on the feedback graphs was presented to participants in terms of "threshold" (the graphs were identical to those used in Study 2 with the exception of altering the labeling of the y-axis). Although in practice this value had no meaning, it had the benefit that it was not obviously capped by a ceiling (i.e., there was no way for a participant to intuit what the maximum level of performance was given that metric). The meaning of "threshold" was explained to the participants simply as "higher scores equal better performance"; we note that while this made the y-axis more opaque to the participant, they should have nonetheless been able to use position along the y-axis to infer whether they were "high" or "low" on the graph. This change then allowed us to include a "high initial skill, high improvement" condition. All told then, the conditions for Study 3 included: "low initial skill, low improvement," "high initial skill, low improvement," "low initial skill, high improvement," "high initial skill, high improvement," and "low initial skill, high improvement (with disjunction)."

Results

Exclusions

No participants were excluded from the analyses conducted for Study 3.

Analyses

Although in Study 3 all comparisons were planned *and* all comparisons of interest were investigated using orthogonal contrasts, for consistency with Study 2 we conducted an omnibus test that included all conditions. This test was significant. $X^2(4) = 35.906$, $p < 0.001$ (noting that this result suggests that we are protected from family-wise Type 1 error rate exceeding 0.05 when performing multiple comparisons).

We next compared our four primary conditions of interest: "low initial skill, low improvement," "high initial skill, low improvement," "low initial skill, high improvement," and "high initial skill, high improvement." As in Study 2, those who never decided to switch tasks were labeled as "censored" (i.e., data were "right" censored). The log-rank test indicated that these groups were significantly different ($p < 0.001$). In other words, the difference in survival probabilities between "low initial skill, low improvement," "high initial skill, low improvement," "low initial skill, high improvement," and "high initial skill, high improvement" conditions was significant. Kaplan–Meier survival curves were estimated for these four groups and are plotted in Fig. 7.

We then compared specific pairs of conditions to investigate the effect of feedback about initial performance (i.e., high or low) and the effect of feedback about performance improvement (i.e., high or low) in a Cox proportional hazards regression analysis. To make these comparisons, we created four orthogonal contrasts (i.e., c1 = c(− 1, − 1,1,1,0), c2 = c(1, − 1,1, − 1,0), c3 = c(− 1,1,1, − 1,0), c4 = c(− 1, − 1, − 1, − 1,4)). Contrast 1 (c1) represented the "initial" comparison: "low initial skill, low improvement" and "low initial skill, high improvement" groups (as a set) were compared to "high initial skill, low improvement" and "high initial skill, high improvement" groups (as a set). To assess the effect of feedback about initial performance, we examined the parameter estimate associated with contrast 1. The direction of the effect was such that the high initial group had a decreased risk of dropping out relative to the low initial group; however, this effect was not significant (Hazard Ratio: 0.75, CI_{95} [0.44,1.27], $p = 0.284$) Additionally, the log-rank test indicated that these groups were not significantly different ($p = 0.57$).

Next, we investigated the effect of feedback about performance improvement on the probability of dropping out (i.e., choosing to switch to a new task). For this analysis, we used the same set of orthogonal contrasts as above, but this time looked at the parameter estimate associated with contrast 2. Contrast 2 (c2) represented the "improvement" comparison: "low initial skill, high improvement" and "high initial skill,

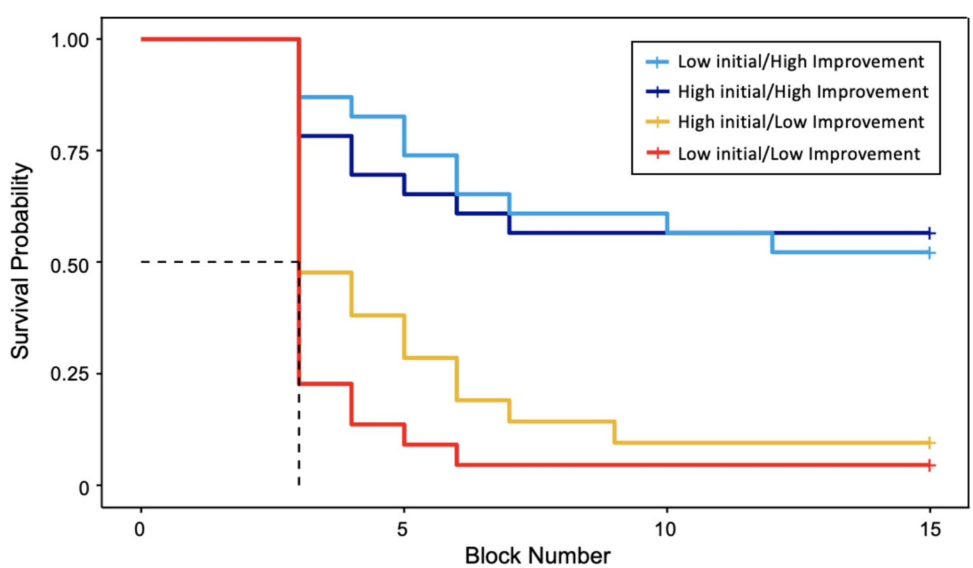

Fig. 7 Kaplan–Meier curves estimated for four primary conditions. The significant log-rank tests suggests that the survival probabilities of these four conditions are different from one another. Furthermore, while the "initial" manipulation (low or high) appeared to make no difference in persistence, the "improvement" manipulation (low or high) produced marked changes in persistence, with participants seeing feedback indicating their performance was improving persisting longer than those who saw feedback indicating that their performance was not changing

high improvement" groups (as a set) were compared to "low initial skill, low improvement" and "high initial skill, low improvement" groups (as a set). The direction of the effect was such that the low improvement group had an increased risk of dropping out relative to the high improvement group. Specifically, participants in a low improvement condition (i.e., those who were told that they were improving slowly) were 4.80 times more likely to choose to switch tasks than those in a high-rate condition (i.e., those who were told that they were improving rapidly) (Hazard Ratio: 4.80, CI_{95} [2.78, 8.28], $p < 0.001$). Likewise, the results of the log-rank test indicated that these two groups were significantly different ($p < 0.001$).

The exploratory analysis in Study 2 (i.e., "low initial skill, high improvement without disjunction" vs "low initial skill, high improvement with disjunction" comparison) prompted a planned analysis in Study 3. We estimated a Cox proportional hazard regression model to assess the difference between the "low initial skill, high improvement (without disjunction)" and "low initial skill, high improvement (with disjunction)" conditions. Results from Study 2 indicated that early disjunctions may lead to greater risk of task switching; thus, we expected to observe the same trend in Study 3.

We created four new orthogonal contrasts to make the desired comparison (i.e., c1 = c(0, − 1,1,1, − 1), c2 = c(0,0, − 1,1,0), c3 = c3 = c(0, − 1,0,0,1), c4 = c(4, − 1, − 1, − 1, − 1)). Contrast 3 (c3) represented the comparison of interest. As expected, the "low initial skill, high improvement (with disjunction)" group was at an increased risk of switching tasks relative to the "low initial skill, high improvement (without disjunction)" group, and this effect was significant (Hazard Ratio: 2.32, CI_{95}[1.09, 4.94], $p = 0.029$). Additionally, the log-rank test indicated that the survival function for these groups was significantly different ($p = 0.032$). Kaplan–Meier curves were estimated for this comparison and are plotted in Fig. 8.

Discussion

As in Study 2, we saw strong support for the hypothesis that manipulations of perceived rate of learning impacted motivation to persist in a cognitive training style task. Participants who were given evidence suggesting that they were improving through time continued with the task far longer than participants who were given evidence that their performance was stagnant. Interestingly, and again mirroring the results of Study 2, our data provided essentially no support for the hypothesis that manipulations of perceived overall performance impacted persistence behavior. Finally, we saw that while participants who were given feedback indicating that they were rapidly improving showed greater persistence than those who were given feedback indicating no improvement, the calculations utilized by participants appear to be quite local. As such, when a minor disjunction downward was observed in an overall clearly increasing performance graph, this had the direct effect of substantially decreasing persistence.

Conclusions and Future Directions

The results here provide a number of directions for future work focused on increasing sustained motivated effort in long-term cognitive training studies. First, at least in the context of the training task utilized here (N-back), participants did not have initially good internal estimates of their own abilities in the absence of feedback. However, they quickly utilized external feedback to construct their view of

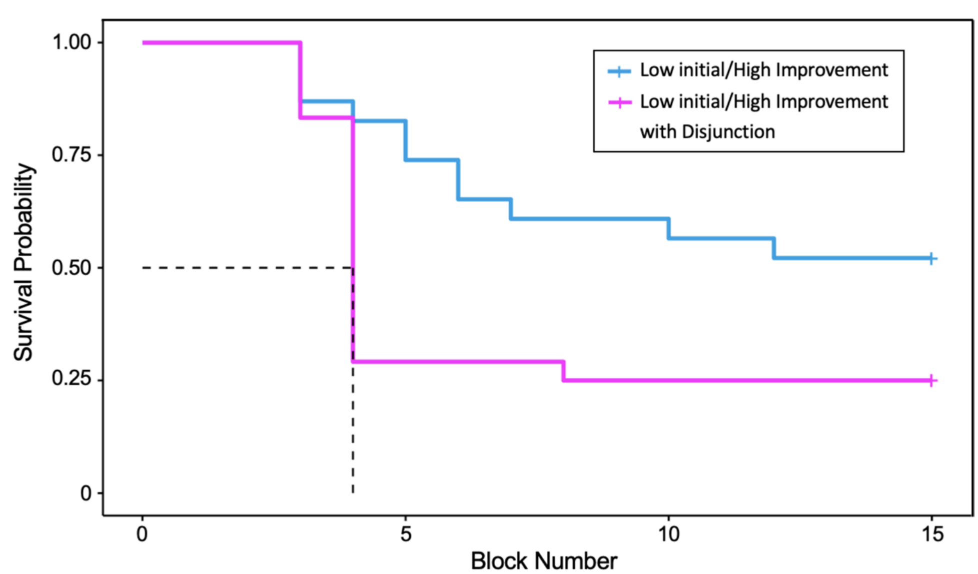

Fig. 8 Kaplan–Meier curves estimated for low initial-high improvement and low initial-high improvement with disjunction groups. Those in the disjunction condition were more likely to switch to a new task than those whose improvement feedback only ever increased. As is clear from the graph, the primary time point that participants in the "with disjunction" group chose to quit was perfectly aligned with the disjunction itself (i.e., as soon as they saw a datapoint indicating that their performance had become "worse" they quit and switched to the other task)

their own performance. This suggests that the feedback that is received, to some extent, sets "reality" for the participants, thus opening potential avenues to purposefully manipulate that feedback to alter persistence behavior. Utilizing more indirect or difficult-to-understand metrics of performance could further augment this effect. Indeed, this is often already seen in metrics in some products (e.g., "Brain Age" Kitahara et al., 2006) wherein the exact links between task performance and feedback are not fully transparent.

Second, participants were more likely to disengage with training when the feedback indicated they were not improving. This presents several challenges for designing feedback. One issue that must be overcome is the tendency for true improvement to appear to flatten out through time. Indeed, the growth in performance that occurs with training often follows a roughly exponential time course (Cochrane & Green, 2021; Cochrane et al., 2024; Kattner et al. 2017b). Such a time course will naturally, at some point, appear to look like a "plateau" or "asymptote" (i.e., even if someone is consistently improving by 50% of the remaining space for improvement, those jumps will start to be smaller and smaller in the raw space as performance improves overall) (Gray, 2017). This might suggest the need to plot feedback in a way that will reduce this tendency (e.g., via some transform). A second issue that must be considered is the tendency for participants to, at some points, show performance that is "worse" than it was in the previous session. If we think of performance in each session as arising from some sampling distribution, it will almost certainly be the case, even if the participant is truly improving monotonically with time, that their performance score in at least one or two sessions will be "worse" than it was in the previous session (with this probability increasing as the full curve becomes shallower). How (or whether) to present such "fallbacks" in performance may require a delicate hand given the risk of participants quitting.

Third, Eccles' expectancy-value theory posits that individuals are more likely to engage in a task when they both value and believe that they can succeed at the task (Eccles & Wigfield, 2020; Eccles et al., 1983; Wigfield & Eccles, 2000). In providing false feedback to participants (Studies 2 and 3), our primary aim was to manipulate participants' beliefs about their ability to succeed at the task. The tendency for participants to quit when feedback indicated they were not improving—even if their performance feedback was positive—suggests that the value they perceived in the task primarily hinged on their competence at the task. Thus, when their competence did not appear to be improving, they saw little reason to persist, even if they believed they could continue to succeed. Promoting participants' *intrinsic* value for the task—i.e., their interest in enjoyment of engaging in the task for its own sake (Eccles & Wigfield, 2020; Eccles et al., 1983; Wigfield & Eccles, 2000), which does not necessarily depend on task competence—may thus be necessary to promote participants' persistence with the task once their abilities have plateaued. Future research that manipulates intrinsic value for cognitive training tasks could shed light on this possibility.

In this vein, a critical next step is to consider how the perceived value of cognitive training programs may be enhanced in the real-world (as opposed to in lab-based settings where it may be difficult to produce a substantial amount of intrinsic value in a task that participants will be engaging in for only a very short time in an obviously experimental setting). Indeed, outside of the lab, individuals are unlikely to even begin engaging in cognitive training of their own volition without first being convinced of the value of such training. This is a particularly difficult problem to address given that the threshold at which a task may be considered valuable enough to influence behavior is likely to differ substantially from individual to individual. One possible route to influencing perceived value might be to link performance on a given task to an improvement in daily life. For example, individuals may be told that technical problem-solving skills—those which might be useful in a wide variety of different occupations—will be enhanced because of progress in training. However, this comes with the knock-on problem that individuals are likely to become demotivated if they fail to observe such effects. It might be most effective, then, to provide value-related information to individuals in a manner that is difficult for participants to directly perceive or that takes advantage of placebo effects. For example, if participants are told that improvements in a cognitive training task lead to improvements in mood or more positive affect, they may experience such effects merely due to expectation. Again, one can see hints of this in some commercial products (e.g., participants may be happy and motivated to continue without perceptible signs of improvement if they're seeing that their "brain age" is at a level that is pleasing).

The research also suggests a number of other future directions. First and foremost, one major limitation of the current work is the short time scale of the training; thus, equivalent work scaled up to more realistic cognitive training durations is necessary. For instance, a single instance of a "disjunction" may produce less quitting behavior in the context of longer training (for any number of reasons, from better estimates of the learning curve to justification-of-effort type phenomena). Second, while falsified feedback could be useful in maximizing motivated effort, feedback often serves multiple purposes in cognitive training designs, with one such purpose being corrective in nature. If the direct links between feedback and performance are altered in order to increase motivated effort, this could have deleterious consequences for learning. If such negative consequences for learning are seen, it would be necessary to determine the appropriate tradeoff point between maximizing motivated effort and maximizing learning. One possibility might be to provide multiple metrics of feedback (with at least one being veridical with performance). Third, our results suggest there may be value in exploring the extent to which individuals are using local derivatives for their engagement decisions in many areas of life, whether it is learning to play a guitar or whether to continue with physical therapy or stay on a diet.

We also acknowledge limits on generalizability with respect to the population sampled in our studies. Participants were college-aged, predominantly white, and from a largely WEIRD demographic background. Middle-aged or older adults as well as individuals from a more diverse range of cultural, educational, and socioeconomic backgrounds may be differentially motivated by feedback about their performance on a cognitive training task for a variety of reasons. In all, continuing to examine how to maximize motivated effort in cognitive training will be key if the paradigms are to do the maximum amount of real-world good—particularly in populations that might not be otherwise motivated to expend long-term effort on training.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s41465-025-00332-6.

Acknowledgements We also acknowledge Tyler Ballweg for methodological design and data collection.

Author Contribution LEA: analyses, manuscript and figure preparation.

AC: methodological design, analyses, manuscript, and figure preparation.

VK: methodological design, data collection.

CAH: manuscript preparation.

CSG: methodological design, manuscript and figure preparation.

Data Availability All data and materials will be made available upon request.

Declarations

Ethics Approval and Consent to Participate All study activities were approved by the University of Wisconsin-Madison Institutional Review Board. All participants gave informed consent.

Consent for Publication Not applicable.

Competing interests One of the authors of this paper is Editor-in-Chief of the journal. This paper was handled by a Senior Editorial Board member who assumed responsibility for its peer review. The Editor-in-Chief was not involved in the peer review process. The other authors have no competing interests to declare that are relevant to the content of this article.

References

Bediou, B., Rodgers, M. A., Tipton, E., Mayer, R. E., Green, C. S., & Bavelier, D. (2023). Effects of action video game play on cognitive skills: A meta-analysis. *Technology, Mind, and Behavior*, *4*(1). https://doi.org/10.1037/TMB0000102

Bland, J. M., & Altman, D. G. (2004). The logrank test. *BMJ: British Medical Journal, 328*(7447), 1073.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Cochrane, A., & Green, C. S. (2021). Assessing the functions underlying learning using by-trial and by-participant models: Evidence from two visual perceptual learning paradigms. *Journal of Vision, 21*(13), 5. https://doi.org/10.1167/jov.21.13.5

Cochrane, A., Lu, Z. L., & Green, C. S. (2024). Modeling perceptual learning as a continuous function of time-on-task increases theoretical specificity and statistical power. *Journal of Cognitive Enhancement, 8*, 394–307. https://doi.org/10.1007/s41465-024-00303-3

Deveau, J., Jaeggi, S. M., Zordan, V., Phung, C., & Seitz, A. R. (2015). How to build better memory training games. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00243

Deveau, J., & Seitz, A. R. (2014). Applying perceptual learning to achieve practical changes in vision. *Frontiers in Psychology, 5*, 1166. https://doi.org/10.3389/fpsyg.2014.01166

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In *Achievement and Achievement Motivation* (pp. 75–146). W. H. Freeman and Company.

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*. https://doi.org/10.1016/j.cedpsych.2020.101859

Gray, W. D. (2017). Plateaus and asymptotes: Spurious and real limits in human performance. *Current Directions in Psychological Science, 26*(1), 59–67. https://doi.org/10.1177/0963721416672904

Green, S. C., Bavelier, D., Kramer, A. F., Vinogradov, S., Ansorge, U., Ball, K. K., Bingel, U., Chein, J. M., Colzato, L. S., Edwards, J. D., Facoetti, A., Gazzaley, A., Gathercole, S. E., Ghisletta, P., Gori, S., Granic, I., Hillman, C. H., Hommel, B., Jaeggi, S. M., … Witt, C. M. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement*, *3*(1), 2–29. https://doi.org/10.1007/s41465-018-0115-y

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging, 23*(4), 692–701. https://doi.org/10.1037/a0014345

Green, C. S., & Bavelier, D. (2015). Action video game training for cognitive enhancement. *Current Opinion in Behavioral Sciences, 4*, 103–108. https://doi.org/10.1016/j.cobeha.2015.04.012

Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., Cambria, J., Flunger, B., Dicke, A.-L., Häfner, I., Brisson, B., & Trautwein, U. (2016). Probing the unique contributions of self-concept, task values, and their interactions using multiple value facets and multiple academic outcomes. *AERA Open, 2*(1), 233285841562688. https://doi.org/10.1177/2332858415626884

Guye, S., Röcke, C., Mérillat, S., von Bastian, C. C., spsampsps Martin, M. (2021). Cognitive training across the adult lifespan. In T. Strobach spsampsps J. Karbach (Eds.), *Cognitive Training: An Overview of Features and Applications* (pp. 141–152). Springer International Publishing. https://doi.org/10.1007/978-3-030-39292-5_10

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100*, 105–122. https://doi.org/10.1037/0022-0663.100.1.105

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509–527.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833.

Kattner, F., Cox, C. R., & Green, C. S. (2016). Transfer in rule-based category learning depends on the training task. *PLoS ONE, 11*(10). https://doi.org/10.1371/journal.pone.0165260

Kattner, F., Cochrane, A., Cox, C., Gorman, T., & Green, C. (2017). Perceptual learning generalization from sequential perceptual training as a change in learning rate. *Current Biology*, *27*. https://doi.org/10.1016/j.cub.2017.01.046

Kattner, F., Cochrane, A., & Green, C. S. (2017b). Trial-dependent psychometric functions accounting for perceptual learning in 2-AFC discrimination tasks. *Journal of Vision, 17*(11), 3. https://doi.org/10.1167/17.11.3

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00242

Katz, B., Jones, M. R., Shah, P., Buschkuehl, M., & Jaeggi, S. M. (2021). Individual differences in cognitive training research. In T. Strobach & J. Karbach (Eds.), *Cognitive Training: An Overview of Features and Applications* (pp. 107–123). Springer International Publishing. https://doi.org/10.1007/978-3-030-39292-5_8

Kitahara, S., Katsuki, Y., & Ito, J. (2006). *Brain age: Train your brain in minutes a day!* [Nintendo DS]. Nintendo.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception, 36*(14), 1–16.

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the "×" out of expectancy-value theory?: A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*, 1058–1066. https://doi.org/10.1177/0956797611415540

Pasqualotto, A., Parong, J., Green, C. S., & Bavelier, D. (2023). Video game design for learning to learn. *International Journal of Human-Computer Interaction, 39*(11), 2211–2228. https://doi.org/10.1080/10447318.2022.2110684

Ruxton, G. D., & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral Ecology, 19*(3), 690–693. https://doi.org/10.1093/beheco/arn020

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience, 2*, 27. https://doi.org/10.3389/fnagi.2010.00027

Schubert, T., Strobach, T., & Karbach, J. (2014). New directions in cognitive training: On methods, transfer, and application. *Psychological Research Psychologische Forschung, 78*(6), 749–755. https://doi.org/10.1007/s00426-014-0619-8

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17*(3), 103–186. https://doi.org/10.1177/1529100616661983

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*(1), 70–83. https://doi.org/10.1037/0012-1649.42.1.70

Vermeir, J. F., White, M. J., Johnson, D., Crombez, G., & Van Ryckeghem, D. M. L. (2020). The effects of gamification on computerized cognitive training: Systematic review and meta-analysis. *JMIR Serious Games, 8*(3), e18644. https://doi.org/10.2196/18644

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology, 48*, 1643–1657. https://doi.org/10.1037/a0027247

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015