

1. 给本文拟定一个abstract (30分) ;
  2. 讨论本研究的贡献、意义和不足之处 (30分) ;
  3. 讨论未来可能的研究方向 (40分) 。
- (第一题用英文作答, 后两题中英文作答皆可)

There is a great deal of current scientific interest as to whether and/or how basic cognitive skills can be improved via dedicated behavioral training (1–3). This potential, if realized, could lead to substantial real-world impact. Indeed, effective training paradigms would have significant value not only for populations that show deficits in cognitive skills (e.g., individuals diagnosed with Attention Deficit Hyperactivity Disorder [ADHD] or Alzheimer’s disease and related dementias) but also, for the general public, where core cognitive capacities underpin success in both academic and professional contexts (4–6). These possible translational applications, paired with an emerging understanding of how to best unlock neuroplastic change across the life span (7, 8), have spurred hundreds of behavioral intervention studies over the past few decades. While the results have not been uniformly positive (perhaps not surprising given the massive heterogeneity in theoretical approach, methods, etc.), multiple meta-analyses suggest that it is possible for cognitive functions to be improved via some forms of dedicated behavioral training (9–11). However, while these basic science results provide optimism that real-world gains could be realized [and in fact, real-world gain is already being realized in some spheres, such as a Food and Drug Administration (FDA)–cleared video game–based treatment supplement for ADHD (12, 13)], concerns have been raised as to whether those interventions that have produced positive outcomes are truly working via the proposed mechanisms or through other nonspecific third-variable mechanisms. Several factors have been proposed to explain improvements in behavioral interventions, including selective attrition, contextual factors, regression to the mean, and practice effects to name a few (14). Here, we focus on whether expectation-based (i.e., placebo) mechanisms can explain improvements in cognitive training (15–17).

In other domains, such as in clinical trials in the pharmaceutical domain for instance, expectation-based mechanisms are typically controlled for by making the experimental treatment and the control treatment perceptually indistinguishable (e.g., both might be clear fluids in an intravenous bag or a white unmarked pill). Because perceptual characteristics cannot be used to infer condition, this methodology is meant to ensure that expectations are matched between the experimental and control groups (both in terms of the expectations that the participants have and in terms of the expectations that the research team members who interact with the participants have). Under ideal circumstances, the use of such a “double-unaware” design ensures that expectations cannot be an explanatory mechanism underlying any differences between the groups’ outcomes [note that we use the double-unaware terminology in lieu of the more common “double-blind” terminology, which can be seen as ableist (18)].

It is unclear whether most pharmaceutical trials do, in fact, truly meet the double-unaware standard (e.g., despite being perceptually identical, active and control treatments nonetheless often produce different patterns of side effects that could be used to infer condition) (19, 20). Yet, meeting the double-unaware standard is particularly difficult in the case of cognitive training interventions (16). Here, there is simply no way to make the experimental and control interventions perceptually indistinguishable while at the same time, ensuring that the experimental condition contains an “active ingredient” that the control condition lacks. In behavioral interventions, no matter what the active ingredient may be, it will necessarily produce a difference in look and feel as compared with a training condition that lacks the ingredient.

Researchers designing cognitive training trials, therefore, typically attempt to utilize experimental and control conditions that, while differing in the proposed active ingredient, will nonetheless produce similar expectations about the likely outcomes (16, 21–24). This type of matching process, however, is inherently difficult as it is not always clear what expectations will be induced by a given type of experience. Consistent with this, there is reason to believe that expectations have not always been successfully matched. In multiple cases, despite attempts to match expectations across conditions, participants in behavioral intervention studies have nonetheless indicated the belief that the true active training task will produce more cognitive gains than the control task (25–27). Critically, the data as to whether differential expectations in these cases actually, in turn, influence the observed outcomes are decidedly mixed. In some cases, participant expectations differed between training and control conditions, and these expectations were at least partially related to differences in behavior (25). In other cases, participants expected to improve but did not show any actual improvements in cognitive skill (28), or the degree to which they improved was unrelated to their stated expectations (29).

Regardless of the mixed nature of the data thus far, there is increasing consensus that training studies should 1) attempt to match the expectations generated by their experimental and control treatment conditions, 2) measure the extent to which this matching is successful and if the matching was not successful, and 3) evaluate the extent to which differential expectations explain differences in outcome (16, 30). Yet, such methods are not ideal with respect to getting to the core question of whether expectation-based mechanisms can, in fact, alter performance on cognitive tasks in the context of cognitive intervention studies in the first place. Indeed, there is a growing body of work suggesting that self-reported expectations do not necessarily fully reflect the types of predictions being generated by the brain (e.g., it is possible to produce placebo analgesia effects even in the absence of self-reported expectation of pain relief) (31, 32). Instead, addressing this question would entail purposefully maximizing the differences in expectations between groups (i.e., rather than attempting to minimize differential expectations and then, measuring the possible impact if the differences were not eliminated, as is done in most cognitive training studies).

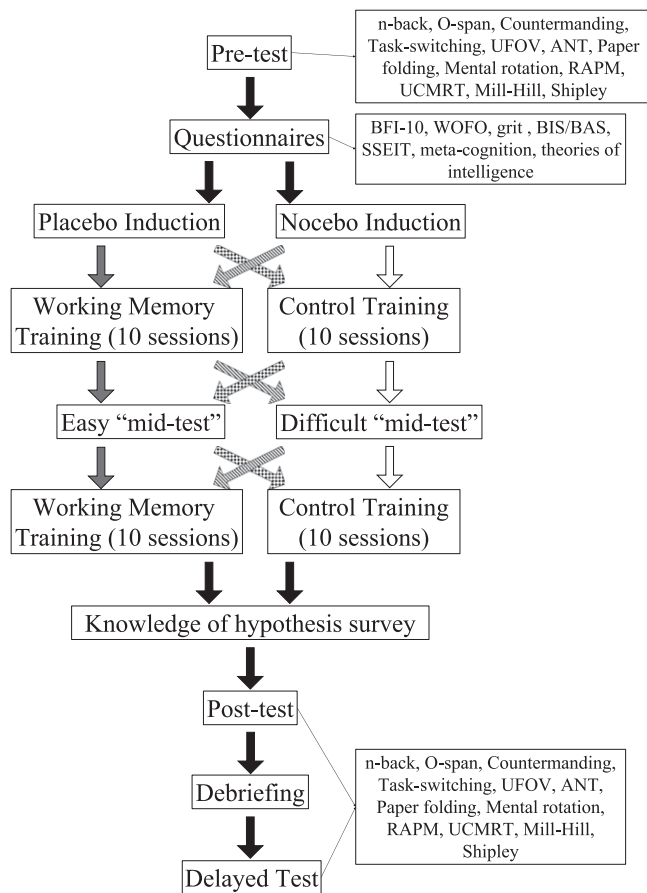
One key question then is how to maximize such expectations. In general, in those domains that have closely examined placebo effects, expectations are typically induced through two broad routes: an explicit route and an associative route. In the explicit route, as given by the name, participants are explicitly told what behavioral changes they should expect (e.g., “this pill will improve your symptoms” or “this cognitive training will improve your cognition”) (33). In the associative learning

route, participants are made to experience a behavioral change associated with expected outcomes (e.g., feeling improvements of symptoms or gains in cognition) through some form of deception (34). For example, in an explicit expectation induction study, participants may first have a hot temperature probe applied to their skin, after which they are asked to rate their pain level. An inert cream is then applied that is explicitly described as an analgesic before the hot temperature probe is reapplied. If participants indicate less pain after the cream is applied, this is taken as evidence of an explicit expectation effect. In the associative expectation version, the study progresses identically as above except that when the hot temperature probe is applied the second time, it is at a physically lower temperature than it was initially (participants are not made aware of this fact). This is meant to create an associative pairing between the cream and a reduction in experienced pain (i.e., not only are they told that the cream will reduce their pain, they are provided “evidence” that the cream works as described). If then, after reapplying the cream and applying the hot temperature probe a third time (this time at the same temperature setting as the first application), if participants indicate even less pain than in the explicit condition, this is taken as evidence of an associative expectation effect. It remains to be clarified how associative learning approaches may be best applied to cognitive training; however, we suggest here that a reasonable approach to this would be to provide test sessions where test items are manipulated to provide participants with an experience where they perceive that they are performing better, or worse in the case of a placebo, than they did at the initial test session. Notably, while there are cases where strong placebo effects have been induced via only explicit (35) or only associative methods (36), in general, the most consistent and robust effects have been induced when a combination of these methods has been utilized (37–39).

Within the cognitive training field, the corresponding literature is quite sparse. Few studies have deliberately attempted to create differences in participant expectations, and of those, all have used the explicit expectation route alone, have implemented the manipulation in the context of rather short interventions (e.g., utilizing 20 min of “training” within a single session rather than the multiple hours that are typically implemented in actual training studies), or both. Of these, the results are again at best mixed, with one study suggesting that expectations alone can result in a positive impact on cognitive measures (40), while others have found no such effects (33, 41, 42). Given this critical gap in knowledge, here we examined the impact of manipulations deliberately designed to maximize the presence of differential expectations in the context of a long-term cognitive training study.

## Current Study

Here, we implemented a 2 (cognitive training vs. active control)  $\times$  2 (positive vs. negative expectation) factorial design to directly estimate the combined effects of explicit and associative placebos. As Fig. 1 shows, the design involved a typical pretest  $\rightarrow$  multisession training  $\rightarrow$  posttest design with two treatment groups (*n*-back working memory training and trivia training); however, it was fully intermixed with two expectation conditions (placebo and placebo). The primary goal of the current study was to examine the extent to which expectation effects can be induced in the context of a typical cognitive training style intervention by focusing on a variety of common targets, including working memory, cognitive flexibility, visual selective attention, spatial cognition, and fluid intelligence. Thus, all



**Fig. 1.** A flowchart of the study procedure and assignment to one of four experimental conditions.

methodological choices were made in the service of maximizing the difference in expectations and thus, the potential magnitude of the expectation effect across groups. For instance, a combination of the explicit and associative routes was utilized to create both the “placebo” (i.e., expectations of improvement as a result of training), and “nocebo” (i.e., expectations of diminished performance as a result of training) conditions. Furthermore, rather than utilizing what current empirical evidence suggests might be the absolute most effective training paradigm in terms of enhancing cognitive function (43), we instead utilized a popular version of *n*-back training that has been shown to have a positive impact in the past but just as importantly, was a better match with our assumptions regarding what participants imagine “brain training” to entail (e.g., more like a psychology task than a video game). This was then paired with an active control task that could also plausibly be expected to improve cognitive function (a general knowledge/trivia task, which takes advantage of naive participants often not distinguishing between crystallized knowledge and fluid cognitive performance) (23). The use of these two training tasks not only ensured that both training conditions could potentially produce an expectation effect, but the use of one task that has more active ingredients than the other (even if it is not what we would consider the strongest possible training task) allowed us to examine whether the presence of these active ingredients interacted with expectations to further augment gains. Next, because in many areas of medicine, there is increasing interest in the use of placebos as treatment, we employed both an immediate posttest at the conclusion of training and a second posttest after participants were fully debriefed regarding the

true purpose of the study. This allowed us to examine not only if expectation effects could be induced but if so, whether these effects were resilient to knowledge about the true nature of the intervention. Previous research has shown that placebo effects persist after participants are made aware of their expectations (44), even if they are not consciously aware of their expectations (45). Finally, the extent to which there are individual difference factors that predict placebo responsiveness has long been of interest to fields that have focused on placebo effects. For example, placebo effects can be predicted by a number of psychological constructs, including goal seeking, self-efficacy, self-esteem, fun or sensation seeking, and neuroticism as well as underlying brain activity (46–49). While this previous research is related, no work to date has examined the role of individual differences in placebo effects in the context of cognitive training. As such, exploratory analyses investigated whether individual characteristics, such as those related to personality and motivation, might predict which types of people are most susceptible to expectation effects in cognitive training (50).

## Results

**Expectation Induction Manipulation Check.** Participants’ stated expectations about how their performance would change after cognitive training aligned with the message they were given prior to training. Those who received the placebo manipulation reported higher expectations than those who received the nocebo manipulation when asked how much cognitive training improves cognition on a scale from 1 (completely unsuccessful) to 7 [completely successful; placebo:  $M = 5.24$ ,  $SD = 1.18$ ; nocebo:  $M = 4.07$ ,  $SD = 1.22$ ;  $t(123) = 5.48$ ,  $P < 0.001$ ], with a 4 on this scale indicating neutral expectations about training. While the nocebo group was intended to create negative expectations (i.e., less than four), previous research has shown that the majority of individuals, even with no explicit expectation given, expect cognitive training to improve cognition (51). Thus, the nocebo manipulation was effective in shifting those expectations to neutral.

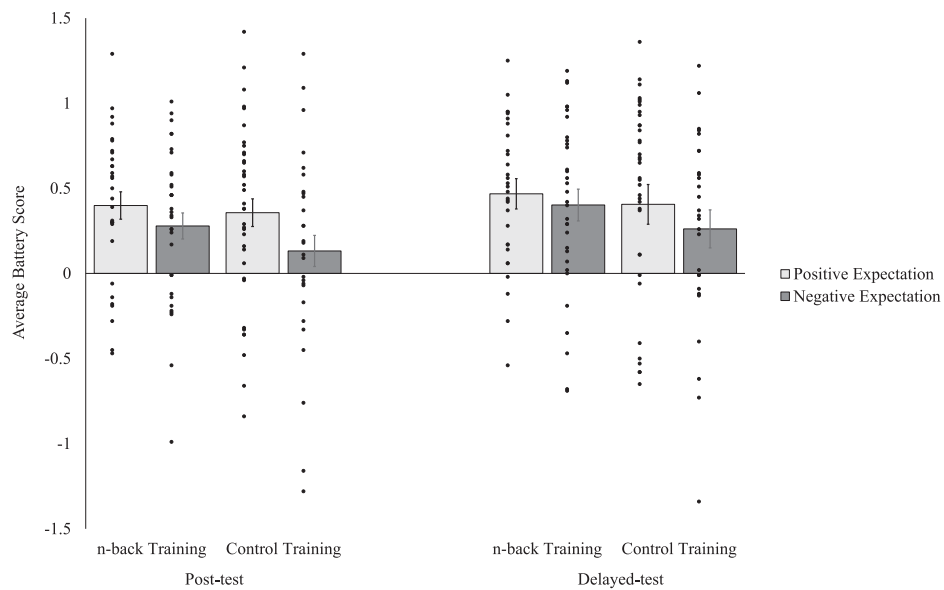
**The Effect of Expectations and Training on Cognitive Battery Performance.** Primary and exploratory analyses followed preregistration on Open Science Framework ( - - - - - ). To determine whether there were any overall expectation or training condition effects on cognitive performance, two 2 (positive vs. nocebo)  $\times$  2 (cognitive training vs. control training) multivariate analyses of covariance (MANCOVAs) were run on the posttest and delayed test task batteries (*n*-back, O span, task-switching task, countermanding tasks, useful field of view [UFOV], attentional network test [ANT], mental rotation task, paper-folding task, University of California Matrix Reasoning Test [UCMRT], and Raven’s Advanced Progressive Matrices [RAPM]) with respective pretest scores as covariates to test several a priori–defined directional hypotheses (i.e., that the placebo condition would out-perform the nocebo condition and that the cognitive training condition would outperform the control condition). Thus, we would reject null hypotheses only if the tests reach the alpha-threshold with respect to these directional hypotheses. At posttest, as shown in Fig. 2, there was a significant effect of the expectation condition, such that those who received the placebo manipulations performed better than those who received the nocebo manipulations [ $F(10,102) = 1.71$ ,  $P = 0.044$ ,  $\eta_p^2 = 0.14$ ]. There was also a significant effect of the training condition, such that those who completed the working memory game performed better than those who completed the control game [ $F(10,102) = 1.89$ ,  $P = 0.027$ ,

$\eta_p^2 = 0.16$ ]. The interaction between expectations and training was also not significant at posttest [ $F(10,102) = 0.82, P = 0.611, \eta_p^2 = 0.07$ ]. At the delayed test, shown in Fig. 2, the effect of the training condition remained significant [ $F(10,102) = 2.47, P = 0.006, \eta_p^2 = 0.20$ ], but the effect of the placebo manipulation no longer showed a significant effect on cognitive performance [ $F(10,102) = 0.99, P = 0.230, \eta_p^2 = 0.09$ ]. The interaction between expectations and training was not significant at delayed test either [ $F(10,102) = 0.38, P = 0.955, \eta_p^2 = 0.04$ ]. Follow-up exploratory analyses were then conducted on a task by task basis. As seen in Table 1, at posttest (again with pretest scores as covariates), those individuals who received the placebo manipulation performed significantly better than those who received the nocebo manipulation on the RAPM, UCMRT, *n*-back, and task-switching task. Training condition, meanwhile, had a significant effect on the *n*-back and the paper-folding task. At the delayed test, as seen in Table 2, the placebo manipulation remained only significant for the task-switching task, and the training condition remained significant for the *n*-back.

**Further Explanation of the Training Effects.** A recently proposed mediation model (52) was tested to determine whether gains on a far-transfer task on a fluid intelligence measure are mediated by near-transfer gains on an untrained *n*-back task, and indeed, we were able to replicate the earlier findings (noting that because this mediation model has only recently been published, a replication of the result was not among our preregistered analyses). In short, the indirect effect of the training group on posttest matrix reasoning performance through *n*-back performance was significant ( $b = -0.18, SE = 0.07, 95\% CI [-0.34, -0.05]$ ), suggesting that the far-transfer gains in the RAPM and UCMRT were due to gains in near-transfer task performance on the *n*-back, thus supporting the previous findings (52).

**Moderator Analyses.** A secondary aim of the study was to examine possible moderating factors of expectation effects on cognitive task performance. While there is a body of work in outside domains examining individual difference factors related to placebo responsiveness, we note from the outset that given the paucity of data in the cognitive domain, these analyses were planned to be exploratory in nature and thus, should be considered as hypothesis generating for future work rather than in any way confirmatory. Because significant expectation effects were found for posttest performance on overall performance as well as specifically on the *n*-back, task-switching, RAPM, and UCMRT tasks, moderator analyses focused on these variables as the dependent variables, with their respective pretest scores as covariates. Moderators examined included gender, age, subscale scores of the Big Five Personality Inventory (openness, conscientiousness, extraversion, agreeableness, and neuroticism), motivation subscale scores from the Behavioral Inhibition System (BIS) and Behavioral Activation System (BAS) scales (drive, fun seeking, reward responsiveness, and BIS total), Grit-scale score, metacognitive-scale score, fixed/growth mindset-scale score, subscale scores of the Schutte Self-Report Emotional Intelligence Test (SSEIT; emotion perception, utilizing emotions, managing self-relevant emotions, and managing others emotions), and subscale scores of the Work and Family Orientation (WOFO) scale (hard work, mastery, and competitiveness). Simple moderations were conducted using Hayes' Process macro for SPSS using 5,000 bootstrap samples for bias correction.

Fig. 3 and Table 3 show the interaction effects between each moderator and the expectation manipulation on overall posttest performance and performance on the *n*-back, task-switching tasks, RAPM, and UCMRT. The effect of the expectation manipulation on *n*-back performance was significantly moderated by extraversion; for those who received the nocebo, *n*-back performance increased as extraversion increased, while for those who received the placebo, *n*-back performance slightly decreased as extraversion increased. Additionally, this effect was also moderated by reward sensitivity, with a similar pattern of results. The effect of the expectation manipulation on task-switching performance was significantly moderated by fun-seeking behavior on the BAS/BIS scale (i.e., the motivation to find novel rewards spontaneously); for those who received the nocebo, task-switching performance increased as fun seeking increased, while for those who received the placebo, task-switching performance decreased as fun seeking increased. The effect of the expectation manipulation on RAPM performance was significantly moderated by mindset; for those who received the nocebo, RAPM performance decreased as growth mindset increased, while for those who received the placebo, RAPM performance slightly increased as growth mindset increased.



**Fig. 2.** Effects of training and expectation conditions on overall performance on the posttest and delayed-test cognitive batteries, controlling for pretest performance. Error bars represent SEM.

**Table 1. Posttest cognitive battery expectation and training effects for all individual tasks included in the MANCOVA with pretest results as a covariate**

Task	Expectation effect			Training effect			Expectation × training interaction		
	<i>F</i>	<i>P</i> value	Partial $\eta^2$	<i>F</i>	<i>P</i> value	Partial $\eta^2$	<i>F</i>	<i>P</i> value	Partial $\eta^2$
<i>n</i> -back	<b>4.20</b>	<b>0.022</b>	<b>0.04<sup>a</sup></b>	<b>12.59</b>	<b>&lt;0.001</b>	<b>0.10</b>	0.03	0.854	0.00
O span	0.17	0.340	0.00	0.21	0.323	0.00	2.16	0.144	0.02
Counter	1.04	0.156	0.01	2.27	0.068	0.02	0.02	0.901	0.00
Task switch	<b>3.41</b>	<b>0.034</b>	<b>0.03</b>	1.07	0.152	0.01	0.28	0.597	0.00
UFOV	0.03	0.428	0.00	0.90	0.172	0.01	0.77	0.382	0.01
ANT	1.00	0.160	0.01	0.02	0.441	0.00	0.47	0.494	0.00
Mental rotation	0.00	0.481	0.00	0.01	0.457	0.00	0.60	0.441	0.01
Paper folding	1.75	0.094	0.02	<b>3.70</b>	<b>0.029</b>	<b>0.03</b>	3.22	0.076	0.03
RAPM	<b>8.39</b>	<b>0.003</b>	<b>0.07</b>	2.25	0.069	0.02	0.57	0.451	0.01
UCMRT	<b>2.93</b>	<b>0.045</b>	<b>0.03</b>	1.43	0.117	0.01	0.27	0.603	0.00

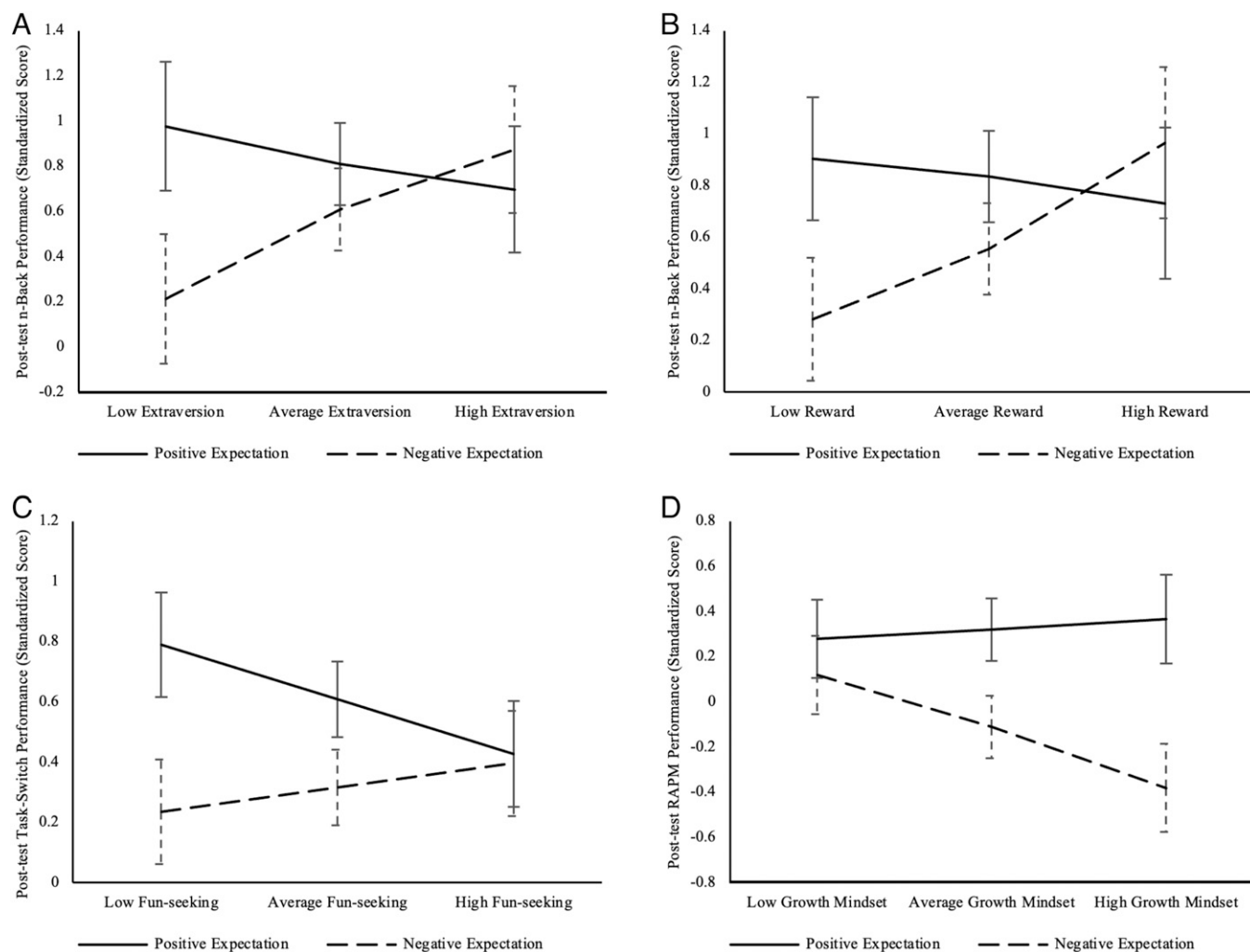
<sup>a</sup>Bold values indicate *p*-values < .05



**Table 2. Delayed-test cognitive battery expectation and training effects for all individual tasks included in the MANCOVA with pretest results as a covariate**

Task	Expectation effect			Training effect			Expectation × training interaction		
	<i>F</i>	<i>P</i> value	Partial $\eta^2$	<i>F</i>	<i>P</i> value	Partial $\eta^2$	<i>F</i>	<i>P</i> value	Partial $\eta^2$
<i>n</i> -back	2.05	0.078	0.02	<b>16.75</b>	<b>&lt;0.001</b>	<b>0.13</b>	0.45	0.502	0.00
O span	0.20	0.329	0.00	0.22	0.320	0.00	0.30	0.586	0.00
Counter	0.36	0.274	0.00	0.09	0.382	0.00	1.10	0.297	0.01
Task switch	<b>4.57</b>	<b>0.018</b>	<b>0.04<sup>a</sup></b>	0.19	0.332	0.00	0.30	0.585	0.00
UFOV	1.58	0.106	0.01	0.02	0.439	0.00	0.26	0.612	0.00
ANT	1.44	0.116	0.01	0.04	0.424	0.00	0.82	0.367	0.01
Mental rotation	0.59	0.223	0.01	0.06	0.407	0.00	0.011	0.744	0.00
Paper folding	0.09	0.384	0.00	2.08	0.076	0.02	0.71	0.406	0.01
RAPM	0.20	0.327	0.00	0.16	0.347	0.00	0.10	0.749	0.00
UCMRT	2.31	0.066	0.02	0.31	0.294	0.00	0.03	0.869	0.00

<sup>a</sup>Bold values indicate *p*-values < .05



**Fig. 3.** Significant moderator analyses for (A) extraversion and the expectation effect on *n*-back performance, (B) the BAS reward subscale and the expectation effect on *n*-back performance, (C) the BAS fun-seeking subscale and the expectation effect on task-switching performance, and (D) mindset and the expectation effect on RAPM performance. All moderation analyses controlled for respective pretest task performance.

**Table 3. Simple moderator analyses for posttest *n*-back, task-switching, RAPM, and UCMRT performance**

Moderator	Overall posttest score			<i>n</i> -back			Task switching			RAPM			UCMRT		
	b	p	CI	b	p	CI	b	p	CI	b	p	CI	b	p	CI
Age	0.02	0.229	-0.02, 0.07	-0.00	0.995	-0.13, 0.13	-0.02	0.737	-0.11, 0.08	-0.01	0.782	-0.12, 0.09	-0.01	0.870	-0.12, 0.10
Gender	0.11	0.367	-0.13, 0.35	0.40	0.316	-0.39, 1.20	0.17	0.549	-0.40, 0.75	-0.17	0.576	-0.79, 0.44	0.20	0.548	-0.46, 0.87
Big 5															
Extra.	0.05	0.075	-0.01, 0.10	<b>0.19</b>	<b>0.035</b>	<b>0.01, 0.36<sup>a</sup></b>	-0.01	0.828	-0.14, 0.11	-0.02	0.825	-0.16, 0.13	-0.12	0.112	-0.27, 0.03
Agree.	0.01	0.772	-0.06, 0.08	-0.03	0.796	-0.26, 0.20	0.06	0.454	-0.10, 0.23	-0.09	0.303	-0.27, 0.09	0.03	0.754	-0.16, 0.22
Consci.	-0.01	0.750	-0.08, 0.06	0.00	0.972	-0.23, 0.23	-0.04	0.674	-0.20, 0.13	-0.05	0.610	-0.23, 0.13	-0.12	0.218	-0.31, 0.07
Neurot.	-0.05	0.142	-0.10, 0.02	0.00	0.980	-0.19, 0.20	0.02	0.786	-0.12, 0.16	0.10	0.202	-0.05, 0.25	-0.04	0.675	-0.20, 0.13
Open.	-0.03	0.301	-0.09, 0.02	0.04	0.659	-0.15, 0.24	0.09	0.204	-0.05, 0.24	0.00	0.970	-0.15, 0.16	-0.04	0.603	-0.21, 0.12
BAS/BIS															
Drive	0.03	0.144	-0.01, 0.08	0.13	0.079	-0.02, 0.27	-0.05	0.382	-0.15, 0.06	0.04	0.462	-0.07, 0.16	-0.06	0.305	-0.19, 0.06
Fun seek	0.02	0.386	-0.03, 0.02	0.06	0.503	-0.11, 0.23	<b>0.13</b>	<b>0.032</b>	<b>0.02, 0.25</b>	0.04	0.579	-0.10, 0.17	0.03	0.716	-0.12, 0.17
Reward	0.01	0.795	-0.04, 0.06	<b>0.17</b>	<b>0.032</b>	<b>0.02, 0.32</b>	-0.00	0.945	-0.12, 0.11	0.01	0.820	-0.11, 0.14	-0.10	0.130	-0.24, 0.03
BIS	-0.01	0.368	-0.04, 0.02	0.06	0.247	-0.04, 0.15	0.01	0.848	-0.06, 0.08	0.03	0.427	-0.05, 0.11	0.01	0.845	-0.07, 0.09
Grit	-0.00	0.656	-0.02, 0.03	0.02	0.533	-0.05, 0.09	-0.04	0.117	-0.09, 0.01	-0.01	0.762	-0.06, 0.05	-0.03	0.318	-0.09, 0.03
Metacog.	0.00	0.226	-0.00, 0.01	0.01	0.251	-0.01, 0.03	-0.00	0.754	-0.02, 0.01	-0.00	0.945	-0.02, 0.02	-0.01	0.407	-0.03, 0.01
Mindset	-0.01	0.057	-0.03, 0.00	-0.01	0.744	-0.06, 0.04	-0.00	0.851	-0.04, 0.03	<b>-0.05</b>	<b>0.018</b>	<b>-0.08, -0.01</b>	-0.04	0.056	-0.08, 0.00
SSEIT															
Percept	-0.01	0.231	-0.03, 0.01	-0.01	0.673	-0.07, 0.04	0.02	0.445	-0.02, 0.05	-0.00	0.887	-0.02, 0.02	-0.01	0.545	-0.06, 0.03
Utility	0.02	0.350	-0.02, 0.05	-0.02	0.624	-0.12, 0.07	-0.03	0.470	-0.10, 0.05	0.00	0.953	-0.08, 0.08	-0.07	0.095	-0.15, 0.01
Self	-0.01	0.231	-0.03, 0.07	0.00	0.927	-0.06, 0.07	-0.01	0.625	-0.06, 0.04	-0.01	0.619	-0.07, 0.04	-0.02	0.406	-0.08, 0.03
Others	0.00	0.954	-0.03, 0.03	0.02	0.651	-0.06, 10	0.05	0.096	-0.01, 0.11	-0.00	0.915	-0.07, 0.06	-0.06	0.062	-0.13, 0.00
WOFO															
Work	0.02	0.334	-0.02, 0.01	0.03	0.666	-0.09, 0.14	-0.02	0.608	-0.11, 0.06	-0.03	0.476	-0.13, 0.06	0.03	0.567	-0.07, 0.13
Mastery	0.02	0.094	-0.00, 0.04	0.03	0.459	-0.05, 0.11	-0.03	0.414	-0.08, 0.03	0.02	0.523	-0.04, 0.08	0.00	0.880	-0.06, 0.07
Compete	-0.01	0.594	-0.04, 0.02	-0.01	0.873	-0.11, 0.09	-0.00	0.983	-0.07, 0.07	-0.03	0.477	-0.11, 0.05	-0.05	0.276	-0.13, 0.04

<sup>a</sup>Bold values indicate p-values < .05. Extra = Extraversion; Agree = Agreeableness; Consci = Conscientiousness; Open = Openness; Metacog = Metacognition

## Materials and Methods

### Method.

**Study overview.** The procedure across all sites involved in the study was approved by the University of Wisconsin–Madison Minimal Risk Research Institutional Review Board, and participants provided written informed consent to participate in the study. Over the first 2 d of the study, participants completed a sizable pretest battery of tasks measuring various cognitive abilities (with multiple measures per cognitive ability) as well as a host of individual-difference questionnaires. Participants were then randomly assigned to their respective expectation conditions (placebo or nocebo) and were given the corresponding explicit expectation information (i.e., that the training they were about to take part in would increase or decrease their performance on the types of tasks they just underwent). They then completed ten 20-min sessions of training across the course of several weeks before completing an associative learning “midtest.” Importantly, this midtest was not meant as an actual evaluation. Instead, it was meant to provide evidence consistent with the explicit expectations that the participants had been initially provided (i.e., that their performance would either improve or worsen). The midtest thus involved the same basic tasks as they had completed at pretest, but the tasks were manipulated to either be easier (placebo) or be more difficult (nocebo) than what was experienced at pretest (with the manipulation being done in such a way that participants could not tell that the tasks themselves had been altered in these ways). Participants then completed an additional ten 20-min sessions of training before returning for a posttest (same 2 d as pretest). Participants were then debriefed as to the true purpose of the study (i.e., were made aware of their assigned expectation and whether it was true relative to their training), after which they completed the cognitive battery one final time. The study was completed completely online via the participants’ personal computers and/or mobile devices, with a researcher present over video conference during all pre-, mid-, and posttesting sessions. The procedure was preregistered on ClinicalTrials.gov.

**Participants.** In this study, 263 participants were initially recruited across three sites located at two West Coast universities ( $n = 94$  and  $n = 104$ ) and one mid-western university ( $n = 94$ ), of which 193 participants ( $n = 60, 84$ , and  $49$ , respectively) completed all sessions. There were 60 participants assigned to the placebo/working memory training condition, 59 assigned to the placebo/control training condition, 62 assigned to the nocebo/working memory training condition, and 54 assigned to the nocebo/control training condition; 57 participants who completed the pretest and were randomly assigned a training and expectation condition dropped out at some point during the study. Critically, the

number of participants who voluntarily dropped from the study at any point ( $n = 57$ ) did not significantly differ by condition [ $\chi^2(3, n = 235) = 3.90, P = 0.273$ ]. Additionally, 28 participants were excluded due to clear noncompliance with the training directions (e.g., did not complete training sessions on time, completed too many), 13 participants had missing data due to computer errors on at least one task, and 27 participants scored outside of three SDs from the mean on at least one task. These participants were excluded from the MANCOVAs as complete data are required for the analysis. In total, 125 participants were included in the final data analyses (mean age = 20.18, range = [13, 37], SD = 2.91). The final sample was 73.6% female, 24.8% male, 0.8% nonbinary, and 0.8% unspecified. Participants were 49.6% Asian, 30.4% White, 9.6% multiracial (two or more races), 0.8% Black or African American, and 9.6% unspecified. Additionally, 77.6% of participants were Hispanic, 21.6% were not Hispanic, and 0.8% did not specify their ethnicity. Participants were recruited using mass emails and flyers targeted to the undergraduate research participant pools at all universities, and thus, the gender/racial/ethnic composition of our sample reflects the student population that typically participates in psychology experiments at the three study sites. Participants were compensated \$170 for completing the study.

### Materials.

**Cognitive task battery: Pretest, posttest, and delayed posttest.** Cognitive skills across multiple cognitive domains, including working memory, cognitive flexibility, visual selective attention, spatial cognition, and fluid intelligence, were measured using 10 tasks.  $z$  scores were first calculated for each cognitive task on the pretest. Standardized scores for each task on the posttest and delayed test were then calculated using the mean and SD from the respective pretest task  $z$ -score distribution [ $(X - \mu_{\text{pretest}})/SD_{\text{pretest}}$ ]. Composite scores for the 10 tasks on the posttest and delayed test were also calculated. Tasks in which the inverse efficiency score (IES) was the dependent variable were reversed coded to ensure that higher scores indicated improved performance for all measures. Two vocabulary measures were included as control measures, as crystallized intelligence should not be influenced by either working memory training or expectation induction (33).

**Working memory tasks.** The  $n$ -back task and a standard complex span task, the O span, were used to measure working memory. In the  $n$ -back task, participants were presented with individual letters sequentially for a brief amount of time (500 ms). Their task was to indicate whether the current letter matched the letter presented  $N$  items back. The  $n$ -back levels varied between two and four back (i.e., target letters matching two, three, or four items previously). Participants completed three blocks of each  $n$ -back level in a random order. Each block contained 15 trials, of which five letters were targets plus  $N$  letters at the beginning of the sequence that could not be targets. The task took about 10 min to complete. The dependent measure was the proportion of hits minus false alarms overall (59, 60). In the O-span task, participants viewed sequences of stimuli that alternated between simple math equations (e.g.,  $[(8 \times 2) - 8 = ?]$ ) and single letters. For each math problem, participants were asked to click the screen after they had mentally solved the problem. They were then presented with a number as the solution to the problem and indicated whether it was true or false. A letter was then briefly presented (1,000 ms). At the end of each sequence of math problem/letter pairs, which ranged between three and seven pairs, they were asked to recall all the letters that they had seen in the presented order. Participants completed three trials of each sequence length in a random order. Participants were asked to make sure their math accuracy remained above 85% correct for the duration of the task. The task took about 14 min to complete. The dependent measure was the sum of all letter sets recalled perfectly (24).

**Cognitive flexibility tasks.** Cognitive flexibility was measured with a task-switching task and a countermanding task. In the task-switching task, participants were shown letter-number pairs (e.g., K7) displayed in one of four quadrants on the screen. The quadrant location indicated to the participants to either categorize the letter as a vowel or consonant (top quadrants) or categorize the number as odd or even (bottom quadrants). The letter-number pairs were presented in a predictable clockwise pattern across each of the four quadrants. Trials in the top right and bottom left quadrants were “nonswitch trials,” in which the participant was asked to perform the same task as on the previous trial, while the trials in the other quadrants were “switch trials,” in which the participant was



asked to perform the opposite task as on the previous trial. Participants completed 72 trials, and the task took about 5 min to complete. The dependent measure was reaction time (RT) and accuracy on switch trials minus nonswitch trials (61). An overall IES ( $IES = RT/\text{proportion of correct trials}$ ) was calculated, with lower IESs indicating better performance. In the countermanding task, participants were presented with two types of stimuli (i.e., a heart or a flower) on either the right or left side of the screen. Their task was to tap one of two buttons on either the same side as the stimulus (e.g., heart; congruent trials) or on the opposite side of the stimulus (e.g., flower; incongruent trials). Participants completed 48 trials, and the task took about 5 min to complete. The dependent variables were RT and accuracy on switch trials minus nonswitch trials (62), and an overall IES was calculated.

**Visual selective attention tasks.** The two visual selective attention measures were the UFOV task and the ANT. In the UFOV task, participants were briefly presented with a central object (car or truck) and then, a display consisting of 48 items on each of the four radial spokes and the four obliques evenly spaced. One of the items located on a radial or oblique was a target (a car), while the remaining items were distractors (black triangles). The participants' task was to indicate which central object was presented and upon which of the eight spokes the target appeared. Participants completed an adaptive number of trials, which terminated after nine reversals in correct/incorrect responses and took about 9 min on average to complete. The dependent measure was the presentation duration threshold (63). In the ANT, participants were first presented with a fixation cross in the center of the screen. Then, one of three cues appeared: no cue, a single-star cue located in the center that indicated the arrow(s) would appear shortly, or a double-star cue located above or below the center that indicated where the arrow(s) would appear. After the cue, an arrow pointing either left or right appeared above or below the center fixation cross. The arrow was flanked on either side by congruent arrows (i.e., arrows facing in the same direction as the center arrow), incongruent arrows (i.e., facing the other direction), or no arrows. The participants' task was to indicate the direction the center arrow was pointing. Participants completed 96 trials, which took about 8 min to complete. The dependent measures were RT and accuracy (64), and an IES was also calculated.

**Spatial cognition tasks.** The two spatial cognition measures were a mental rotation task and a paper-folding task. In the mental rotation task, participants were shown two images side by side. The images were either identical but with one rotated relative to the other or were mirror reversed and rotated copies of one another. The participants' task was to indicate whether the two items were identical or were mirror-reversed copies. Participants completed 36 trials, and the task took about 6 min to complete. The percentage correct was the dependent measure (65). In the paper-folding task, participants were shown a piece of paper in a series of images being folded various ways before a hole was punched in the paper. They were then asked to indicate from five alternative answers what the paper would look like when unfolded. The task consisted of 10 trials, with a maximum of 3 min to complete. The percentage of correctly solved items was the dependent measure (66).

**Fluid intelligence assessments.** The two measures of fluid intelligence were RAPM (67) and UCMRT (68). In both tasks, the participant was presented with a grid of elements with one of the elements missing and was asked to identify the missing element that completes the grid pattern. The RAPM consisted of 14 problems and had a maximum time limit of 9 min. The UCMRT consisted of 16 problems and had a maximum time limit of 8 min. The dependent variable for both tasks was the percentage of correctly solved items.

**Control tasks.** Two vocabulary control tasks were used as control measures: the vocabulary sections of the *Mill Hill Vocabulary Scale* (69) and the *Shipley Institute of Living Scale* (70). These require participants to select the appropriate synonym for a target word among several alternatives. The *Mill Hill Vocabulary Scale* consisted of 24 words, and the *Shipley Institute of Living Scale* consisted of 15 words. The dependent variable for both tasks was the percentage of correct responses.

**Questionnaires.** Participants completed a series of questionnaires to provide demographic variables (e.g., age, gender, race/ethnicity, socioeconomic status) as well as to assess other individual difference factors that may be predictive of placebo responsiveness. Personality traits were measured using the Big Five Inventory - 10 item (BFI-10), which included items assessing conscientiousness, agreeableness, neuroticism, openness to experience, and extraversion (71). The WOFO scale measured achievement motivation and included 19 items in

separate scales for work (positive attitudes toward hard work), mastery (preference for difficult, challenging tasks), and competitiveness (72). The Grit scale was an eight-item scale that measured the tendency to sustain effort toward long-term goals (73). Sensitivity to reward and avoidance was measured using the BIS/BAS (74). Dweck's Intelligence scale measured the extent to which people believe that intelligence is fixed or malleable (55). The SSEIT measured the ability to know and understand one's own emotions and the emotions of others, which included subscales of emotion perception, utilization of emotion, managing self-relevant emotions, and managing others' emotions (75). Lastly, the Meta-Cognitive Skills Scale measured the ability to understand one's own abilities, strengths, and weaknesses (76).

#### **Training tasks.**

**Working memory training.** Participants who were assigned to the true cognitive training condition completed a nongamified version of Recollect the Study, which was based on the *n*-back task. During each training session, participants were sequentially presented with individual colored circles. Their task was to tap the screen when the current circle's color matched the one presented *N* items back. The circles were presented for 2,500 ms (interstimulus interval [ISI] = 500 ms), of which 30% were targets and another 30% were lures (e.g., items that occur *N* - 1 or *N* + 1 of the target position). Each training session was broken into short ~2-min training blocks, with typically 7 to 10 blocks per session. Visual feedback and auditory feedback were presented on all trials, indicating correct and incorrect answers. The task difficulty was adaptive and adjusted the *n*-back level based on the player's performance. Participants completed twenty 20-min sessions over the course of 2 to 4 wk. Participants were allowed to choose which days they completed sessions, with restrictions of completing only one session per day and to have no more than 3 d between consecutive sessions. The outcome measure was average *N* level per day (59).

**Control training.** Participants who were assigned to the control training condition played a trivia game, Knowledge Builders. In this game, participants were presented with multiple-choice questions consisting of general knowledge, vocabulary, social science, and trivia, and it has been used in previous research (23, 33, 59). Like the working memory training, participants also completed twenty 20-min sessions with the same completion requirements. This game was adaptive and adjusted the difficulty of questions based on the performance on the previous level. This task was used as the active control activity as it taps into crystallized intelligence, and performance is not expected to be influenced by working memory training, as previously shown (77).

#### **Expectation inductions.**

**Explicit expectations.** Participants assigned to the placebo group received information that suggested that cognitive training improved cognitive skills, including performance both on the trained task and on the untrained tasks in the study. Conversely, participants assigned to the nocebo group received information that suggested that while cognitive training might improve performance on that trained task, performance on all other untrained tasks would decrease. In both conditions, participants were presented with a short explanation of why the cognitive training is expected to improve or decrease their cognitive performance, which included exaggerated or false scientific studies to support these expectations.

**Associative learning-based expectations.** The midtest consisted of altered tasks from the Cognitive Task Battery except the *n*-back task as it would be expected to interact with the working memory training. Those in the placebo group received tasks that were manipulated to be easier than the original tasks, while those in the nocebo group received tasks that were manipulated to be more difficult than the original tasks. Critically, in all cases, this was done in a way where it would be difficult, if not impossible, for participants to identify the changes. For example, in the mental rotation task, the placebo version was manipulated such that it included more trials with small-angle rotations (e.g., 10° to 90°; which are "easier"—they can be completed faster and more accurately) and fewer trials with large-angle rotations (i.e., 110° to 190°), while the nocebo version included the opposite.

**Knowledge of hypothesis survey.** As a manipulation check, a knowledge of hypothesis survey was administered to determine whether the explicit expectations given to participants indeed created the appropriate expectations about the outcomes of the cognitive training they completed. Participants were asked

to rate how much they thought they knew the researchers' hypotheses and to describe what the specific hypotheses were. Additionally, a similar survey was administered to all research assistants after each session in which they interacted with participants to determine whether researchers also remained unaware of participants' expectations.

1. V. Pergher *et al.*, Divergent research methods limit understanding of working memory training. *J. Cogn. Enhanc.* **4**, 100–120 (2020).
2. P. D. Harvey, S. R. McGurk, H. Mahncke, T. Wykes, Controversies in computerized cognitive training. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **3**, 907–915 (2018).
3. D. Bavelier, C. S. Green, A. Pouget, P. Schrater, Brain plasticity through the life span: Learning to learn and action video games. *Annu. Rev. Neurosci.* **35**, 391–416 (2012).
4. C. S. Green, N. S. Newcombe, Cognitive training: How evidence, controversies, and challenges inform education policy. *Policy Insights Behav. Brain Sci.* **7**, 80–86 (2020).
5. A. Genevsky, C. T. Garrett, P. P. Alexander, S. Vinogradov, Cognitive training in schizophrenia: A neuroscience-based approach. *Dialogues Clin. Neurosci.* **12**, 416–421 (2010).
6. J. A. Anguera, A. Gazzaley, Video games, cognitive exercises, and the enhancement of cognitive abilities. *Curr. Opin. Behav. Sci.* **4**, 160–165 (2015).
7. M. Nahum, H. Lee, M. M. Merzenich, Principles of neuroplasticity-based rehabilitation. *Prog. Brain Res.* **207**, 141–171 (2013).
8. A. Pahor, S. M. Jaeggi, A. R. Seitz, "Brain training" in *eLS* (Mauro MacCarone, Ed., John Wiley & Sons, Ltd., Hoboken, NJ, 2018), pp. 1–9.
9. B. Bediou *et al.*, Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychol. Bull.* **144**, 77–110 (2018).
10. A. M. Tetlow, J. D. Edwards, Systematic literature review and meta-analysis of commercially available computerized cognitive training among older adults. *J. Cogn. Enhanc.* **1**, 559–575 (2017).
11. J. Au *et al.*, Improving fluid intelligence with training on working memory: A meta-analysis. *Psychon. Bull. Rev.* **22**, 366–377 (2015).
12. S. H. Kollins, A. Childress, A. C. Heusser, J. Lutz, Effectiveness of a digital therapeutic as adjunct to treatment with medication in pediatric ADHD. *NPJ Digit. Med.* **4**, 58 (2021).
13. S. W. Evans *et al.*, The efficacy of cognitive videogame training for ADHD and what FDA clearance means for clinicians. *EPCAMH* **6**, 116–130 (2021).
14. T. A. Salthouse, *Major Issues in Cognitive Aging* (Oxford University Press, 2010), vol. **49**.
15. S. Denker *et al.*, Assessing the impact of expectations in cognitive training and beyond. *J. Cogn. Enhanc.* **5**, 502–518 (2021).
16. C. S. Green *et al.*, Improving methodological standards in behavioral interventions for cognitive enhancement. *J. Cogn. Enhanc.* **3**, 2–29 (2019).
17. D. J. Souders *et al.*, Evidence for narrow transfer after short-term cognitive training in older adults. *Front. Aging Neurosci.* **9**, 41 (2017).
18. D. Morris, S. Fraser, R. Wormald, Masking is better than blinding. *BMJ* **334**, 799 (2007).
19. A. Hróbjartsson, E. Forfang, M. T. Haahr, B. Als-Nielsen, S. Brorson, Blinded trials taken to the test: An analysis of randomized clinical trials that report tests for the success of blinding. *Int. J. Epidemiol.* **36**, 654–663 (2007).
20. G. H. Montgomery, I. Kirsch, Classical conditioning and the placebo effect. *Pain* **72**, 107–113 (1997).
21. T. Klingberg *et al.*, Computerized training of working memory in children with ADHD—a randomized, controlled trial. *J. Am. Acad. Child Adolesc. Psychiatry* **44**, 177–186 (2005).
22. Y. Brehmer, H. Westerberg, L. Bäckman, Working-memory training in younger and older adults: Training gains, transfer, and maintenance. *Front. Hum. Neurosci.* **6**, 63 (2012).
23. S. M. Jaeggi, M. Buschkuhl, P. Shah, J. Jonides, The role of individual differences in cognitive training and transfer. *Mem. Cognit.* **42**, 464–480 (2014).
24. S. M. Jaeggi, M. Buschkuhl, J. Jonides, P. Shah, Short- and long-term benefits of cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10081–10086 (2011).
25. P. L. Baniqued *et al.*, Working memory, reasoning, and task switching training: Transfer effects, limitations, and great expectations? *PLoS One* **10**, e0142169 (2015).
26. S. M. Hynes, Internet, home-based cognitive and strategy training with older adults: A study to assess gains to daily life. *Aging Clin. Exp. Res.* **28**, 1003–1008 (2016).
27. S. N. Yeo *et al.*, Effectiveness of a personalized brain-computer interface system for cognitive training in healthy elderly: A randomized controlled trial. *J. Alzheimers Dis.* **66**, 127–138 (2018).
28. S. Guye, C. C. von Bastian, Working memory training in older adults: Bayesian evidence supporting the absence of transfer. *Psychol. Aging* **32**, 732–746 (2017).
29. R.-Y. Zhang *et al.*, Action video game play facilitates "learning to learn." *Commun. Biol.* **4**, 1154 (2021).
30. J. Au, B. C. Gibson, K. Bunarjo, M. Buschkuhl, S. M. Jaeggi, Quantifying the difference between active and passive control groups in cognitive interventions using two meta-analytical approaches. *J. Cogn. Enhanc.* **4**, 192–210 (2020).
31. S. M. Schäfer, L. Colloca, T. D. Wager, Conditioned placebo analgesia persists when subjects know they are receiving a placebo. *J. Pain* **16**, 412–420 (2015).
32. K. B. Jensen *et al.*, A neural mechanism for nonconscious activation of conditioned placebo and nocebo responses. *Cereb. Cortex* **25**, 3903–3910 (2015).
33. N. Tsai *et al.*, (Un)great expectations: The role of placebo effects in cognitive training. *J. Appl. Res. Mem. Cogn.* **7**, 564–573 (2018).
34. M. Vodyanik, A. Cochrane, A. Corriveau, Z. Demko, C. S. Green, No evidence for expectation effects in cognitive training tasks. *J. Cogn. Enhanc.* **5**, 296–310 (2021).
35. L. Koban, M. Jepma, M. López-Solà, T. D. Wager, Different brain networks mediate the effects of social and conditioned expectations on pain. *Nat. Commun.* **10**, 4096 (2019).
36. N. J. Voudouris, C. L. Peck, G. Coleman, Conditioned response models of placebo phenomena: Further support. *Pain* **38**, 109–116 (1989).
37. T. D. Wager, L. Y. Atlas, The neuroscience of placebo effects: Connecting context, learning and health. *Nat. Rev. Neurosci.* **16**, 403–418 (2015).
38. M. Zunhammer, T. Spisák, T. D. Wager, U. Bingel, Placebo Imaging Consortium, Meta-analysis of neural systems underlying placebo analgesia from individual participant fMRI data. *Nat. Commun.* **12**, 1391 (2021).
39. E. Carlino *et al.*, Role of explicit verbal information in conditioned analgesia. *Eur. J. Pain* **19**, 546–553 (2015).

**Data, Materials, and Software Availability.** Cognitive battery results have been deposited in Open Science Framework (78).

**ACKNOWLEDGMENTS.** This research was supported by National Institute on Aging Grants R56AG063952 (to A.R.S., S.M.J., and C.S.G.) and K02AG054665 (to S.M.J.).

40. C. K. Foroughi, S. S. Monfort, M. Paczynski, P. E. McKnight, P. M. Greenwood, Placebo effects in cognitive training. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7470–7474 (2016).
41. S. Rabipour *et al.*, Few effects of a 5-week adaptive computerized cognitive training program in healthy older adults. *J. Cogn. Enhanc.* **4**, 258–273 (2020).
42. B. Katz, S. M. Jaeggi, M. Buschkuhl, P. Shah, J. Jonides, The effect of monetary compensation on cognitive training outcomes. *Learn. Motiv.* **63**, 77–90 (2018).
43. J. Deveau, S. M. Jaeggi, V. Zordan, C. Phung, A. R. Seitz, How to build better memory training games. *Front. Syst. Neurosci.* **8**, 243 (2015).
44. L. Colloca, R. Klinger, H. Flor, U. Bingel, Placebo analgesia: Psychological and neurobiological mechanisms. *Pain* **154**, 511–514 (2013).
45. L. Colloca, F. Benedetti, How prior experience shapes placebo analgesia. *Pain* **124**, 126–133 (2006).
46. K. Weimer, L. Colloca, P. Enck, Placebo effects in psychiatry: Mediators and moderators. *Lancet Psychiatry* **2**, 246–257 (2015).
47. B. Horing, K. Weimer, E. R. Muth, P. Enck, Prediction of placebo responses: A systematic review of the literature. *Front. Psychol.* **5**, 1079 (2014).
48. T. D. Wager, L. Y. Atlas, L. A. Leotti, J. K. Rilling, Predicting individual differences in placebo analgesia: Contributions of brain activity during anticipation and pain experience. *J. Neurosci.* **31**, 439–452 (2011).
49. L. Koban, L. Ruzic, T. D. Wager, "Brain predictors of individual differences in placebo responding" in *Placebo and Pain: From Bench to Bedside*, L. Colloca, M. A. Flaten, K. Meissner, Eds. (Academic Press, 2013), pp. 89–102.
50. A. L. Geers, S. G. Helfer, K. Kosbab, P. E. Weiland, S. J. Landry, Reconsidering the role of personality in placebo effects: Dispositional optimism, situational expectations, and the placebo response. *J. Psychosom. Res.* **58**, 121–127 (2005).
51. S. Rabipour, P. S. R. Davidson, Do you believe in brain training? A questionnaire about expectations of computerised cognitive training. *Behav. Brain Res.* **295**, 64–70 (2015).
52. A. P. Pahor, A. Seitz, S. M. Jaeggi, Getting at underlying mechanisms of far transfer: The mediating role of near transfer. *PsyArXiv [Preprint]* (2021). <https://psyarxiv.com/q5xmp/> (Accessed 30 May 2022).
53. W. R. Boot, D. J. Simons, C. Stothart, C. Stutts, The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspect. Psychol. Sci.* **8**, 445–454 (2013).
54. K. J. Klauer, G. D. P. Hyne, Inductive reasoning: A training approach. *Rev. Educ. Res.* **78**, 85–123 (2008).
55. C. Dweck, *Self-Theories: Their Role in Motivation, Personality and Development* (Taylor and Francis, Philadelphia, PA, 2002).
56. C. L. Collins *et al.*, Video-based remote administration of cognitive assessments and interventions: A comparison with in-lab administration. *J. Cogn. Enhanc.* **6**, 316–326 (2022).
57. J. L. Hardy *et al.*, Enhancing cognitive abilities with comprehensive training: A large, online, randomized, active-controlled trial. *PLoS One* **10**, e0134467 (2015).
58. S. Mohammed *et al.*, The benefits and challenges of implementing motivational features to boost cognitive training outcome. *J. Cogn. Enhanc.* **1**, 491–507 (2017).
59. S. M. Jaeggi *et al.*, The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence* **38**, 625–635 (2010).
60. N. Unsworth, R. P. Heitz, J. C. Schrock, R. W. Engle, An automated version of the operation span task. *Behav. Res. Methods* **37**, 498–505 (2005).
61. R. D. Rogers, S. Monsell, The costs of a predictable switch between simple cognitive tasks. *J. Exp. Psychol. Gen.* **124**, 207–231 (1995).
62. M. C. Davidson, D. Amso, L. C. Anderson, A. Diamond, Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* **44**, 2037–2078 (2006).
63. K. K. Ball, B. L. Beard, D. L. Roenker, R. L. Miller, D. S. Griggs, Age and visual search: Expanding the useful field of view. *J. Opt. Soc. Am. A* **5**, 2210–2219 (1988).
64. J. Fan, B. D. McCandless, T. Sommer, A. Raz, M. I. Posner, Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* **14**, 340–347 (2002).
65. L. Cooper, R. Shepard, "Chronometric studies of the rotation of mental images" in *Visual Information Processing*, W. G. Chase, Ed. (Academic Press, New York, NY, 1973), pp. 135–142.
66. A. Lovett, K. Forbus, Modeling spatial ability in mental rotation and paper-folding. *Proc. Annual Meeting Cogni. Sci. Soc.* **35**, 930–935 (2013).
67. J. Raven, J. Raven, "Raven progressive matrices" in *Handbook of Nonverbal Assessment*, R. S. McCullum, Ed. (Kluwer Academic/Plenum Publishers, 2003), pp. 223–237.
68. A. Pahor, T. Stavropoulos, S. M. Jaeggi, A. R. Seitz, Validation of a matrix reasoning task for mobile devices. *Behav. Res. Methods* **51**, 2256–2267 (2019).
69. J. Raven, J. Raven, R. Court, *Mill Hill Vocabulary Scale* (Oxford University Press, Oxford, United Kingdom, 1998).
70. W. C. Shipley, *Shipley Institute of Living Scale* (Western Psychological Services, Los Angeles, CA, 1986).
71. B. Rammstedt, O. P. John, Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *J. Res. Pers.* **41**, 203–212 (2007).
72. J. Adams, R. F. Priest, H. T. Prince, Achievement motive: Analyzing the validity of the WOFO. *Psychol. Women Q.* **9**, 357–369 (1985).
73. A. L. Duckworth, C. Peterson, M. D. Matthews, D. R. Kelly, Grit: Perseverance and passion for long-term goals. *J. Pers. Soc. Psychol.* **92**, 1087–1101 (2007).
74. N. G. Poythress *et al.*, Psychometric properties of Carver and White's (1994) BIS/BAS scales in a large sample of offenders. *Pers. Individ. Dif.* **45**, 732–737 (2008).
75. X. Gong, S. E. Paulson, Validation of the Schutte self-report emotional intelligence scale with American college students. *J. Psychoeduc. Assess.* **36**, 175–181 (2018).
76. S. M. Fleming, H. C. Lau, How to measure metacognition. *Front. Hum. Neurosci.* **8**, 443 (2014).
77. S. M. Jaeggi *et al.*, Investigating the effects of spacing on working memory training outcome: A randomized, controlled, multisite trial in older adults. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **75**, 1181–1192 (2020).
78. J. Parong, A. R. Seitz, S. M. Jaeggi, C. S. Green, Expectation effects in working memory training. Open Science Framework. <https://osf.io/bder4/files/osfstorage>. Deposited 10 May 2022.