aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
8(4) 445–454

DOI: 10.1177/1745691613491271
pps.sagepub.com
SAGE

The Pervasive Problem With Placebos in Psychology: Why Active Control Groups Are Not Sufficient to Rule Out Placebo Effects

Walter R. Boot[1], Daniel J. Simons[2], Cary Stothart[1], and Cassie Stutts[1]

[1]Department of Psychology, Florida State University and [2]Psychology Department & Beckman Institute for Advanced, Science and Technology, University of Illinois at Urbana-Champaign

Abstract

To draw causal conclusions about the efficacy of a psychological intervention, researchers must compare the treatment condition with a control group that accounts for improvements caused by factors other than the treatment. Using an active control helps to control for the possibility that improvement by the experimental group resulted from a placebo effect. Although active control groups are superior to "no-contact" controls, only when the active control group has the same expectation of improvement as the experimental group can we attribute differential improvements to the potency of the treatment. Despite the need to match expectations between treatment and control groups, almost no psychological interventions do so. This failure to control for expectations is not a minor omission—it is a fundamental design flaw that potentially undermines any causal inference. We illustrate these principles with a detailed example from the video-game-training literature showing how the use of an active control group does not eliminate expectation differences. The problem permeates other interventions as well, including those targeting mental health, cognition, and educational achievement. Fortunately, measuring expectations and adopting alternative experimental designs makes it possible to control for placebo effects, thereby increasing confidence in the causal efficacy of psychological interventions.

Keywords

intervention design, research methods, placebo effect, demand characteristics

To draw causal conclusions about the efficacy of a psychological intervention, researchers must compare the treatment condition with a baseline or control group that accounts for improvements caused by factors other than the treatment. In pharmacological research, the control group receives a sugar pill (a placebo) that looks identical to the experimental pill, meaning that participants cannot tell whether they are in the experimental condition or the control condition. Because they are blind to their condition assignment, they should not hold different expectations for the effectiveness of the pill, and any difference between the groups on the outcome measure may be attributed to the effect of the treatment.[1]

Compared with drug trials, psychological interventions face bigger challenges in accounting for placebo effects. Participants in psychological interventions typically know which treatment they received. For example, participants undergoing an experimental cognitive therapy for anxiety are aware that they are receiving treatment and are likely to expect to improve as a result. Measuring the effectiveness of this therapy by comparing it with a no-treatment control condition would be inadequate because the two groups would have different expectations for improvement, and few scientists would accept such a comparison as compelling evidence that

Corresponding Author:
Walter R. Boot, Department of Psychology, Florida State University, 1107 W. Call Street, Tallahassee, FL 32306–4301
E-mail: boot@psy.fsu.edu

the ingredients of the therapy were responsible for observed improvements. A better comparison would be with an *active* control group, one that receives a similar therapy that does not specifically target their anxiety.

Many researchers, reviewers, and editors of psychology interventions apparently believe that including an active control group automatically controls for placebo effects. We have come to this conclusion because published papers regularly include causal claims about the effectiveness of an intervention without any attempt to test whether the experimental and control groups shared the same expectations. This failure to control for the confounding effect of differential expectations is not a minor omission—it is a fundamental design flaw that potentially undermines any causal inference. Absent any measurement of expectations, conclusions about the effectiveness of an intervention, whether the intervention is designed to improve education, mental health, well-being, or perceptual and cognitive abilities, are suspect. We should distrust those conclusions just as we discount findings from a drug study in which participants knew they were getting the treatment.

To illustrate how such a lack of verification undermines claims of intervention effectiveness, we examine in detail the claim that action video-game training enhances perceptual and cognitive abilities. We focus on the game-training literature not because it is a particularly egregious example of poor design, but because it is better than most—unlike many other psychology interventions, game-training studies typically include active control conditions that are closely matched to the training condition. Nevertheless, they still do not adequately account for expectation effects.

A Case Study: Do Action Video Games Improve Cognition?

We studied the relationship between expectation effects and actual improvement by measuring expectation of improvement directly in two survey studies and comparing our results to the literature on the effects of action video-game interventions. Critically, we measured how expectations differ between the experimental and control conditions in such interventions. We then evaluate the concordance between intervention effects and expectation effects and discuss the implications for understanding action-game effects.

Performance

In many previous training studies, participants who trained for 10 to 50 hr on fast-paced, visually demanding action video games showed improved performance on a variety of perceptual and cognitive measures that tap visual processing, attention, and task-switching (e.g., Green & Bavelier, 2003, 2006a, 2006b, 2007; Green, Sugarman, Medford, Klobusicky, & Bavelier, 2012; Li, Polat, Makous, & Bavelier, 2009; Li, Polat, Scalzo, & Bavelier, 2010; Strobach, Frensch, & Schubert, 2012; but see also Boot, Blakely, & Simons, 2011; Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Kristjánsson, 2013). Most video-game training studies compare improvements for an action-game group with those for an active control group that played a slower-paced, nonaction game (e.g., *Tetris* or *The Sims*) for an equivalent amount of time (e.g., Green & Bavelier, 2003, 2006a, 2006b, 2007; Green et al., 2012). However, no study has tested whether participants trained on slow-paced games such as *The Sims* or simple games such as *Tetris* expect to see improvements on cognitive and perceptual tasks. More precisely, participants in these control groups might not expect the same amount of improvement on the same tasks as do participants playing fast-paced, visually demanding action games (first-person shooters) like *Medal of Honor* and *Unreal Tournament* (Boot et al., 2011; Boot & Simons, 2012).

Measuring expectations

We explicitly measured expectations for improvement in two survey studies (200 participants each). Participants first watched a short video of either an action game (*Unreal Tournament*) or one of the commonly used control games (*Tetris* or *The Sims*). Next, they learned about a set of cognitive and perceptual tasks often used as outcome measures in such studies. For each, they read a description of the task, viewed a video showing what a participant would see when performing the task, and indicated whether they thought their performance on that task would improve as a result of training on the video game they had viewed earlier. If the control game conditions (*Tetris, The Sims*) are an adequate placebo control for the action game-training condition (*Unreal Tournament*), participants viewing the control games should expect the same levels of improvement on each outcome measure as those viewing the action game.

Comparing expectations and actual improvement

Survey respondents viewing an action game expected greater improvement in the same tasks that actually show greater improvements in an intervention study. For example, after action-game training, participants show improved performance on vision and attention tasks, including the useful field of view (UFOV) and multiple object tracking (MOT). Participants who viewed the action game, *Unreal Tournament*, were significantly more likely to believe that training would improve UFOV

performance than were those who viewed *Tetris*, and participants who viewed *Unreal Tournament* were more likely to believe training would improve both UFOV performance and MOT performance than those who viewed *The Sims* (see Table 1 for data and statistics). Note that those viewing an action game did not expect greater improvements on all outcome measures—their expectations were task-specific. For example, they did not expect greater performance on a story-memory task compared with participants who viewed *Tetris* or *The Sims*.[2] These survey results also have implications for claims that *Tetris* training can improve spatial skills (Boot et al., 2008; De Lisi & Wolford, 2002; Okagaki & Frensch, 1994; Sims & Mayer, 2002). Participants who viewed *Tetris* were more likely to believe that training would improve mental rotation performance than were those who viewed *Unreal Tournament*. Consistent with this difference, the only study to directly compare training on an action game with training on *Tetris* found improved mental rotation performance only for the *Tetris* group (Boot et al., 2008).

Table 1. Correspondence Between Differential Expectations in Our Surveys and the Results of Training Studies

Task	Greater Expectation	Greater Improvement
MOT	Action Game *Unreal* (86%) > *Sims* (65%) χ^2 (1, N = 200) = 11.92, p < .01	Action game[a]
UFOV	Action Game *Unreal* (83%) > *Tetris* (45%) χ^2 (1, N = 200) = 31.34, p < .01 *Unreal* (75%) > *Sims* (38%) χ^2 (1, N = 200) = 27.85, p < .01	Action game[b]
MR	Nonaction Game *Tetris* (89%) > *Unreal* (69%) χ^2 (1, N = 200) = 12.06, p < .01	Nonaction game[c]
SM	None, Nonaction Game *Unreal* (37%) = *Tetris* (28%) χ^2 (1, N = 200) = 1.85, p = .17 *Sims* (51%) > *Unreal* (35%) χ^2 (1, N = 200) = 5.22, p = .02	Untested

Note: Data represent the percentage of participants who believed the game they viewed would improve performance on a specific task. Participants viewed a video of video game play (*Unreal Tournament, Tetris*, or *The Sims*), and then viewed videos of cognitive tasks with a description of each task (MOT, UFOV, MR, SM). Participants were then asked to judge whether training on the game they viewed would improve the performance of each cognitive measure. Each video was approximately 30 s long. Data were collected through Amazon's Mechanical Turk system. Participants were 18 years of age or older, living in the United States, paid $0.20, and randomly assigned a game video to view, with the order of cognitive-task videos randomly determined. MOT = multiple object tracking, UFOV = useful field of view, MR = mental rotation, SM = story memory.
[a]Green & Bavelier (2006b). [b]Green & Bavelier (2003, 2006a). [c]Boot et al. (2008).

Implications for training studies

The pattern of expected results reported by untrained participants consistently fit the published results of game-training interventions. Consequently, greater improvements by an action-game group than a control group in an intervention study do not justify the conclusion that game training improves cognition. Not only have the studies failed to control for placebo effects, but also our surveys suggest that such differential expectations for improvement are likely to be present in gaming interventions. The pattern of expectations was comparable regardless of whether participants were familiar with media coverage of the benefits of game training and whether participants reported being gamers themselves.

Our participants formed expectations about the possible benefits of a game after only 30 s of exposure. Presumably, extensive game exposure (10–50 hours of game play), coupled with feedback about gaming improvements, could induce even stronger expectations. Moreover, in our study, participants viewed only one game, so they could not compare their "intervention" with the other one. It is likely that differential expectations would be even greater if participants were aware not only of their own intervention but also of the other one (e.g., if both groups complete their training in the same lab rooms at the same time).

These results underscore the need for experimenters to eliminate (or at least test for) differential expectations. No published study has done so, and the pattern of results in our surveys confirms that such differential placebo effects are entirely plausible as an explanation for all published claims of benefits from gaming interventions. Consequently, the active control conditions used as a baseline comparison for action-game training do not permit causal conclusions about the efficacy of game training. Researchers must demonstrate the absence of placebo effects before concluding in favor of the presence of an intervention effect (see Fig. 1 for a flow-chart illustrating what we can conclude from an intervention).

A Broader Problem

Although our example singled out video-game interventions, the placebo problem is pernicious and pervasive, affecting most cognitive interventions in psychology. For example, one highly cited intervention study (Mahncke

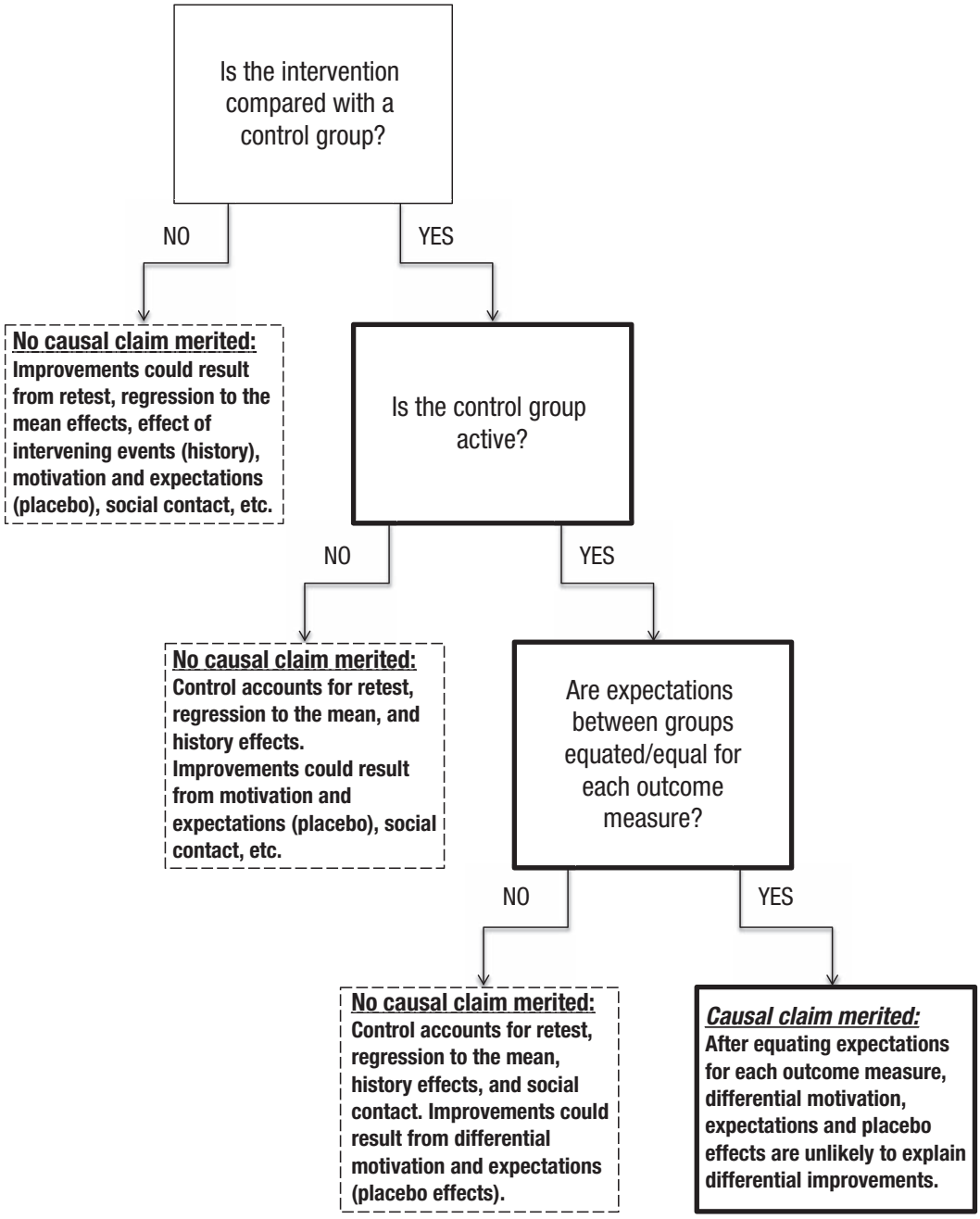

Fig. 1. Appropriate conclusions from a study in which the experimental group improves from pretest to posttest.

et al., 2006) compared training with a commercial brain-fitness program with two control groups: an active group and a no-contact control group. The training group completed auditory tasks that adapted to participants' performance, continuously challenging them. The active control group watched educational DVDs and performed only the pretest and posttest tasks (i.e., their learning from the DVDs was not tested). Compared with the two control groups, the brain fitness group improved more from pretraining to posttraining on a different set of auditory memory tasks.

This finding has been used to promote the scientific effectiveness of a commercial brain-fitness training program, but it lacks an adequate control for placebo effects, meaning that it does not provide compelling evidence for the effectiveness of the intervention. First, the similarity of the training tasks to the outcome measures means that the training group probably would have a greater reason

to expect improvements; participants who watched DVDs have little reason to expect improved auditory memory performance. The authors took the lack of a difference between the DVD group and the no-contact control as evidence ". . . that there is no meaningful placebo effect." This inference is premature. The active control group provided no check against a differential placebo effect because it did not equate the expectations to those of the intervention group. Remarkably, the authors concluded that lack of difference between the DVD and no-contact control groups means that "future studies may not need to include both types of control groups." Dispensing with active control groups altogether would invalidate any conclusions about training effectiveness. Only with an appropriate active control group, one that equates expectations to those of the training group, can an intervention draw a causal conclusion about training effectiveness.

As another example, take the exciting claim that adaptive memory exercises improve IQ in both children and adults. Most studies have included only a no-contact control, which does not eliminate placebo effects (e.g., Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Rudebeck, Bor, Ormond, O'Reilly, & Lee, 2012). In fact, when other researchers did measure expectations (Redick et al., 2012), those who received memory training believed that they had shown improved intelligence, memory, and ability to complete daily activities after training. Similar to our video-game results, in the absence of an active control group that equates for expected performance improvements on each outcome measure, any actual improvements might be explained by perceived or expected benefits rather than actual benefits.

Devising an appropriate control condition for a psychological intervention can be challenging. Take the case of the link between playing violent games and aggression (e.g., Anderson et al., 2010; Ferguson & Kilburn, 2009). Participants viewing graphic materials or playing violent games in a lab are likely to expect a link to aggression, or at least more of a link than they would for playing nonviolent puzzle, sports, or racing games. What active control condition could overcome the surface plausibility of that association, thereby eliminating expectation effects and demand characteristics (Ferguson & Dyck, 2012; see Adachi & Willoughby, 2011, for an alternative explanation for violent game effects when games are more closely matched)? Whenever the effect of an intervention maps onto participant beliefs about what *should* result from an intervention, definitive claims about the effect of the intervention itself are inappropriate.

These placebo problems are not limited to cognitive interventions. Take the claim that daily writing improves physical and mental health (see Pennebaker, 1996, for review). In such studies, participants in the experimental group typically write (repeatedly) about personal thoughts and feelings, experienced trauma, or highly emotional issues. In contrast, those in the control condition typically write about trivial topics (e.g., "Describe the outfit you are wearing today in detail" or "Describe the things you do before class on a typical Monday"; Park & Blumberg, 2002). Matching the activity in the experimental and the active control group is laudable, but the two groups presumably differ in their expectations for therapeutic benefits, meaning that any improvements might result from a differential placebo effect.

The lack of placebo controls in psychotherapy interventions is not new, and it has been discussed for decades (e.g., Rosenthal & Frank, 1956). But it persists. Consider the newly emerging field of internet-based psychotherapy. Many studies use only waitlist controls, some use control conditions in which participants simply read information about their condition, and others use online support groups with no guidance or interaction with an online therapist (e.g., Carlbring et al., 2011; B. Klein, Richards, & Austin, 2006). Without a control for differential expectations, the mechanisms through which these interventions produce their effect (placebo or nonplacebo) are difficult to know.

A Way Forward

The lack of masked condition assignment in psychological interventions is not a minor inconvenience—it is a fundamental design flaw, and experimenters have an obligation to test for the possible consequences of these design limitations. Although some have claimed that placebo control groups in psychological interventions, such as ones examining the effect of game play on cognition, are impossible (Bavelier & Davidson, 2013), that limitation does not excuse researchers from the requirement to account for expectation effects before inferring that an intervention was effective. There are methods to measure and account for the influence of differential expectations and demand characteristics. These include explicitly assessing expectations, carefully choosing outcome measures that are not influenced by differential expectations, and using alternative designs that manipulate and measure expectation effects directly.

Assessing expectations

Our surveys illustrate one approach that can test for the possibility of differential placebo effects in already-published intervention studies. Using Amazon Mechanical Turk, we found that the active control conditions typically used in video-game studies provide an inadequate baseline because participants believe that the action-game treatments will produce bigger improvements in visual processing than will the control games. The same

approach could be used for other interventions, both as a check for placebo problems and as a way to choose outcome measures for future interventions.

For example, participants undergoing an aerobic exercise intervention show greater cognitive improvements than do those in stretching and toning control groups (e.g., Colcombe & Kramer, 2003; Kramer et al., 1999). By recruiting a separate group of participants, describing each intervention (or having them participate in one or two sessions), and checking their expectations, it would be possible to test whether differential expectations are consistent with the pattern of training benefits. Although this method might not generate expectations as strong as engaging in the entire intervention, it could be one of several checks on expectations, and it could help when selecting the most appropriate active control task. In the case of exercise and cognition, we suspect the pattern of expectations would be comparable for the treatment and control conditions, but without empirical verification, differential placebo effects are a possibility. Again, the lack of such tests is not a minor omission—such checks are a necessary precondition for causal claims given the lack of a truly double-blind design.

In addition to checking for differential expectations after the fact, researchers could test for them during the study itself (e.g., O. Klein et al., 2012; Orne, 1969). This method has the advantage that expectation and improvement can be measured in the same subjects (the danger, though, is that tests of expectancy may be reactive). As a hypothetical example, consider a driving intervention aimed at reducing reaction time to road hazards in a sample of older drivers. If participants in each condition were asked to report their beliefs in the efficacy of the training, the pattern shown in Figure 2a would be comforting: Participants' beliefs are not systematically related to the degree of improvement. The pattern in Figure 2b would be cause for concern, though: It is consistent with an effect driven by expectations rather than the treatment.[3] See Serfaty, Csipke, Haworth, Murad, and King (2011) for a careful consideration of potential expectation effects in the depression literature and Redick et al. (2012) in the cognitive training literature.

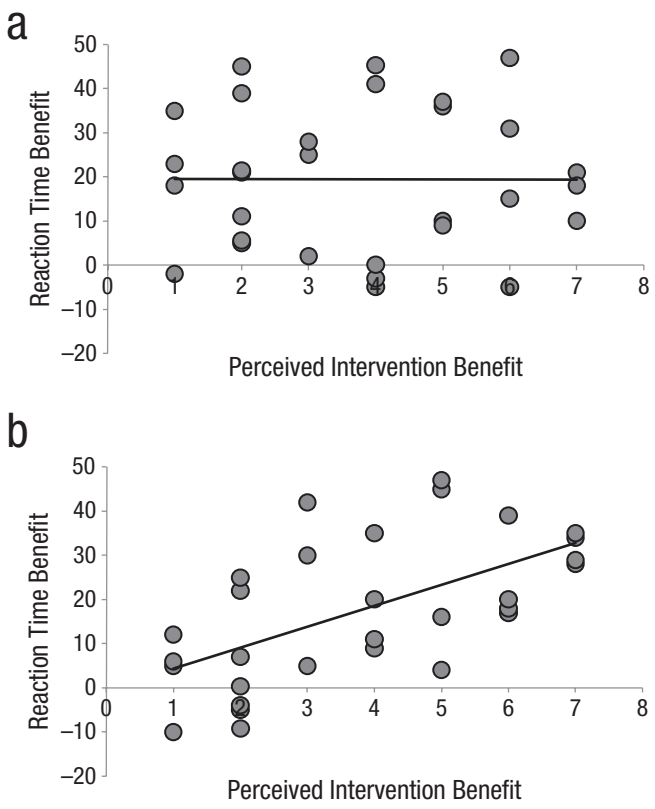

Fig. 2. Graphs from hypothetical data showing reaction time benefit as a function of perceived intervention benefit. In Panel a, improvements in response time are unrelated to an individual's expectation for improvement. This pattern provides evidence that expectations did not drive improvements in performance. In Panel b, improvements in response time were positively related to expected improvements. This pattern suggests the possibility of a placebo effect and potentially undermines any claim about the effectiveness of an intervention.

Choosing the right tasks

Even better than measuring expectations during a study or after the fact would be to choose an active control task or outcome measure on the basis of an independent assessments of expectations. For example, a game-training study could choose an outcome measure that shows no difference in expectations between the action game and control game but that the hypothesis predicts should benefit from action-game training. An even stronger manipulation would choose an outcome measure in which participants expect to benefit more from the control game. If training on the action game then produced greater improvements, the effect could not result from differential expectations.

Note that differential expectations do not necessarily account for differential improvements; such expectations might not have causal potency either, and differential expectations might not produce differences in actual performance across conditions. However, the presence of differential expectations undermines claims about the power of a treatment. Only by isolating the active ingredient of the experimental treatment can we draw firm causal conclusions about its impact.

One possible way to isolate a treatment effect from differential expectations is to demonstrate, empirically, that expectations cannot influence performance on an outcome measure. A task that is objectively impervious to experimentally increased motivation or expectations should be less subject to placebo effects in a training study. For example, if task performance is unchanged by

giving a large incentive for good performance, then different expectations for improvement on that task might have little effect. Such a null effect of motivation on task performance provides a check on the causal potency of differential expectations.

Researchers could take this approach one step further by maximizing motivation to perform well on the pretraining tasks. If subjects are highly motivated and incentivized to perform well during the pretest, then any further improvements are less likely to result solely from expectations of improvement. This procedure would provide a better baseline to isolate the effect of the treatment. As we note later, however, expectations can have effects that go beyond increasing motivation to perform well.

Alternative designs

When it is ethical, experimenters could manipulate expectations directly to test whether a particular outcome measure is sensitive to expectation effects (O'Leary & Borkovec, 1978). For example, in a *neutral expectancy* design, half of each group (experimental and active control) is led to believe that the intervention they are receiving will improve their outcome, whereas the other half is led to have neutral expectations (see Clifasefi, Takarangi, & Bergman, 2006, for an example in the alcohol intoxication literature). In a *counterdemand* design, participants are led to believe that benefits will accrue only after a specified amount of training or experience, and they are tested before and after this period. By directly manipulating expectations, these designs help isolate the effects of expectation from other effects of an intervention.

A dose-response design, in which different groups receive different amounts of treatment, might also be diagnostic; a cognitive-training intervention that produces the same degree of effect on an outcome measure after one training session as after 100 training sessions is suspect. However, dose-response effects could still result from changing expectations as a function of the amount of treatment experienced. *Component control* manipulations, to some extent, also address the effect of expectancy on outcomes. In this method, a multicomponent intervention serves as the experimental treatment, whereas the same treatment minus one component serves as the control. Given the similarity of each treatment, placebo effects are less likely (although researchers still must test for them). Such designs help isolate the possible mechanisms responsible for improvement (for an example of this method in the video-game and cognition literature, see Brown, May, Nyman, & Palmer, 2012). However, if the active control group still contains enough of the active ingredient, then it might show benefits as well. Although *component control* designs provide specificity about possible causal mechanisms underlying improvements, they do not necessarily eliminate differential expectations.

"Just a Placebo Effect?"

We have discussed placebo effects largely in terms of expectations influencing the motivation to perform well on an outcome measure (e.g., someone devoting more effort to a memory measure after completing memory training because he or she now expects to perform better). However, placebo effects can operate in other ways and take many forms (for review, see Benedetti, Mayberg, Wager, Stohler, & Zubieta, 2005; Price, Finniss, & Benedetti, 2008).

Much of the work on the power of placebo effects has focused on pain reduction. Placebos can trigger the release of endogenous opioids and can also reduce pain through nonopioid mechanisms (Montgomery & Kirsch, 1996). Placebo treatments are associated with functional brain changes, including decreased activity in pain-related brain areas (Wager et al., 2004). Placebos also can operate via classical conditioning: If the act of taking medication is associated with a physiological response, an inert placebo can trigger a similar conditioned response (Stockhorst, Steingrüber, & Scherbaum, 2000). Finally, expectancies can affect memory for previous experiences (Price et al., 1999), biasing self-report and subjective outcome measures in favor of an intervention.

Placebo effects are real and worthy of explanation in their own right, and we do not mean to dismiss their important (and clinically relevant) effects in medical and psychological interventions. However, whenever researchers want to attribute causal potency to the intervention itself, it is incumbent on them to verify that the improvements are not driven by expectations.

Setting the Bar Too High?

Given the challenges inherent in conducting psychology interventions, studies necessarily lack some of the critical controls of a double-blind clinical trial. Even studies with weak control conditions can provide useful speculative evidence for possible causal relationships, though, particularly early in a field's development. Although expectations can and should be assessed in all intervention studies, when they are not, researchers should temper causal conclusions appropriately and discuss potential placebo effects explicitly.

Is it unfair to demand adequate testing of and control for placebo effects in all psychological interventions? We think not, but others may disagree. Below we address several of the more common reactions to these guidelines that we have encountered in our discussions with colleagues and in the literature.

The requirement to control for placebo problems will make it too difficult to "get an effect"

In other words, imposing a requirement for adequate active control conditions will produce too many false negatives in studies of training benefits (Schubert & Strobach, 2012). Balancing the risk of missing a real effect against the risk of false positives is essential. However, those risks must be considered in light of the consequences of not knowing whether effects are due to the treatment itself or to participants' expectations. We do not see why controlling for the confound of differential expectations undermines the chances of finding a true benefit if one exists.

The early, exploratory stages of research should tolerate less rigorous adherence to methodological standards

Perhaps the initial study in a field should have license to use less-than-ideal control conditions to identify possible treatments if the authors acknowledge those limits. Even then, a study lacking appropriate controls risks wasting effort, money, and time as researchers pursue false leads. Moreover, the methods of an initial, flawed study can become entrenched as standard practice, leading to their perpetuation; new studies justify their lack of control by citing previous studies that did the same. For that reason, we argue that any intervention, even one addressing a new experimental question, should include adequate tests for expectation effects.

Our methods are better than those used in other psychology intervention studies

All intervention studies should use adequate controls for placebo effects, and the fact that other studies neglect such controls does not justify substandard practices. For example, the use of active control conditions in the video-game-training literature is better than the common use of no-contact controls in the working-memory-training literature, but that does not excuse the lack of placebo controls in either. "Everyone else is doing it" does not justify the use of a poor design.

Converging evidence overcomes the weaknesses in any individual study, thereby justifying causal conclusions

Replication and converging evidence are welcome, but convergence means little if individual studies do not eliminate confounds. In some areas, such as the video-game literature, researchers often appeal to cross-sectional data comparing gamers with nongamers as converging evidence that games cause changes in perception and cognition. Of course nonexperimental studies suffer from a host of other problems (namely third variable and directionality problems), and such designs do not permit any causal conclusions (Boot et al., 2011; Kristjánsson, 2013). Converging evidence is useful in bolstering causal claims only to the extent that we have confidence in the methods of the individual studies providing the evidence.

Final Thoughts

Expectation effects and placebo effects are known problems and, in many ways, are interesting in their own right. In some cases, whether improvements result from the treatment or from a placebo effect is irrelevant; if the expectation that a treatment will alleviate anxiety leads to less anxiety, the patient still benefits (although demand characteristics may be more of a concern, in this case leading to "benefits" that appear only in the laboratory). Many treatments in use today might work in part through placebo effects or work better through an interaction between nonplacebo and placebo effects. However, because we are scientists interested in *mechanisms* of improvement, and our research is funded on the basis of understanding the causal efficacy of treatments, it matters whether improvements are placebo-driven. Only when we know the mechanisms through which improvements occur can we design interventions that tap those mechanisms.

Despite full awareness of the reasons for and benefits of double-blind designs, psychologists persist in drawing inappropriate inferences from designs that lack adequate controls. Without measuring and controlling for placebo effects, such studies provide little more than speculation about the causes of improvements. In the case of cognitive interventions, the field has had enough speculation. Researchers, reviewers, and editors should no longer accept inadequate control conditions, and causal claims should be rejected unless a study demonstrably eliminates differential placebo effects. We are hopeful that, with better designs and better checks on placebo effects, future research will provide more compelling evidence for the effectiveness of interventions. We have outlined a number of methods, designs, and approaches that, when considered together, can lead to a better understanding of how psychological interventions induce improvements.

Acknowledgments

D. J. Simons and W. R. Boot designed the survey and developed the idea for the article. D. J. Simons implemented the surveys using materials prepared by C. Stutts and C. Stothart. C. Stothart conducted the statistical analyses, and W. R. Boot

and C. Stutts prepared figures. W. R. Boot and C. Stothart wrote the first draft of the manuscript, and D. J. Simons and W. R. Boot edited and revised it. Reported data are available online at http://www.opensciencefframework.org/project/7EB6A/

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. An exception would be if the experimental drug induced noticeable side effects. If so, participants might become unblinded to their condition, requiring additional checks to ensure differential expectations were not responsible for observed changes on outcome measures (e.g., see Moscucci, Byrne, Weintraub, & Cox, 1987).
2. Expectations for improvement are not the same as motivation to perform well in general, and controlling for overall motivation is not the same as controlling for differential placebo effects. Differences in overall motivation might lead to improvements on all tasks, but differential expectations might lead to improvements that are more selective. Some papers conflate these two effects, arguing that controlling for overall motivation eliminates placebo effects. For example, Cain, Landau, and Shimamura (2012) showed that action-game players did not outperform non–action-game players on a story-memory task and used that lack of a difference to argue against a placebo effect or demand characteristics as an explanation for group differences on other tasks. Our survey shows why that inference is unmerited. Our participants also expected no differential improvement on the story-memory task, but they showed different expectations for other tasks. The concern about inadequate control groups applies both to differences in overall motivation and to differential expectations for individual tasks as a function of the training condition.
3. Note that it also might be due to awareness of the actual effectiveness of the intervention. The results do not distinguish between these possibilities. A slope difference between the experimental and control conditions would suggest two different mechanisms for improvement, one driven by expectations and one by the intervention.

References

Adachi, P. J., & Willoughby, T. (2011). The effect of video game competition and violence on aggressive behavior: Which characteristic has the greatest influence? *Psychology of Violence*, *1*, 259–274.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., . . . Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries. *Psychological Bulletin*, *136*, 151–173.

Bavelier, D., & Davidson, R. J. (2013). Brain training: Games to do you good. *Nature*, *494*, 425–426.

Benedetti, F., Mayberg, H. S., Wager, T. D., Stohler, C. S., & Zubieta, J. K. (2005). Neurobiological mechanisms of the placebo effect. *Journal of Neuroscience*, *25*, 10390–10402.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology*, *2*, 226.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*, 387–398.

Boot, W. R., & Simons, D. J. (2012). Advances in video game methods and reporting practices (but still room for improvement): A commentary on Strobach, Frensch, and Schubert (2012). *Acta Psychologica*, *141*, 276–277. doi:10.1016/j.actpsy.2012.06.011

Brown, C., May, R., Nyman, J., & Palmer, E. (2012). Effects of friend vs. foe discrimination training in action video games. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting: Vol. 56* (pp. 1416–1420). Thousand Oaks, CA: SAGE Press.

Cain, M. S., Landau, A. N., & Shimamura, A. P. (2012). Action video game experience reduces the cost of switching tasks. *Attention, Perception, & Psychophysics*, *74*, 641–647. doi:10.3758/s13414-012-0284-1

Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthan, E., . . . Andersson, G. (2011). Individually-tailored, Internet-based treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy*, *49*, 18–24.

Clifasefi, S. L., Takarangi, M. K., & Bergman, J. S. (2006). Blind drunk: The effects of alcohol on inattentional blindness. *Applied Cognitive Psychology*, *20*, 697–704.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, *14*, 125–130.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *Journal of Genetic Psychology*, *163*, 272–282.

Ferguson, C. J., & Dyck, D. (2012). Paradigm change in aggression research: The time has come to retire the General Aggression Model. *Aggression and Violent Behavior*, *17*, 220–228.

Ferguson, C. J., & Kilburn, J. (2009). The public health risks of media violence: A meta-analytic review. *Journal of Pediatrics*, *154*, 759–763.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537. doi:10.1038/nature01647

Green, C. S., & Bavelier, D. (2006a). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1465–1478. doi:10.1037/0096-1523.32.6.1465

Green, C. S., & Bavelier, D. (2006b). Enumeration versus object tracking: Insights from video game players. *Cognition*, *101*, 217–245.

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, *18*, 88–94. doi:10.1111/j.1467-9280.2007.01853.x

Green, C. S., Sugarman, M. A., Medford, K., Klobusicky, E., & Bavelier, D. (2012). The effect of action video game

experience on task-switching. *Computers in Human Behavior*, *28*, 984–994. doi:10.1016/j.chb.2011.12.020

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, USA*, *105*, 6829–6833.

Klein, B., Richards, J. C., & Austin, D. W. (2006). Efficacy of internet therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *37*, 213–238.

Klein, O., Doyen, S., Leys, C., Magalhães de Saldanha da Gama, P. A., Miller, S., Questienne, L., & Cleeremans, A. (2012). Low hopes, high expectations: Expectancy effects and the replicability of behavioral experiments. *Perspectives on Psychological Science*, *7*, 572–584. doi:10.1177/1745691612463704

Kramer, A. F., Hahn, S., Cohen, N., Banich, M., McAuley, E., Harrison, C., . . . Colcombe, A. (1999). Aging, fitness, and neurocognitive function. *Nature*, *400*, 418–419.

Kristjánsson, A. (2013). The case for causal influences of action videogame play upon vision and attention. *Attention, Perception, & Psychophysics*, *75*, 667–672.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game playing. *Nature Neuroscience*, *12*, 594–551.

Li, R., Polat, U., Scalzo, F., & Bavelier, D. (2010). Reducing backward masking through action game training. *Journal of Vision*, *10*, 33. doi:10.1167/10.14.33

Mahncke, H. W., Connor, B. B., Appelman, J., Ahsanuddin, O. N., Hardy, J. L., Wood, R. A., . . . Merzenich, M. M. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: A randomized, controlled study. *Proceedings of the National Academy of Sciences, USA*, *103*, 12523–12528.

Montgomery, G., & Kirsch, I. (1996). Mechanisms of placebo pain reduction: An empirical investigation. *Psychological Science*, *7*, 174–176.

Moscucci, M., Byrne, L., Weintraub, M., & Cox, C. (1987). Blinding, unblinding, and the placebo effect: An analysis of patients' guesses of treatment assignment in a double-blind clinical trial. *Clinical Pharmacology & Therapeutics*, *41*, 259–265.

Okagaki, L. R., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescents. *Journal of Applied Developmental Psychology*, *15*, 33–58.

O'Leary, K. D., & Borkovec, T. D. (1978). Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *American Psychologist*, *33*, 821–830.

Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143–179). New York: Academic Press.

Park, C. L., & Blumberg, C. J. (2002). Disclosing trauma through writing: Testing the meaning-making hypothesis. *Cognitive Therapy and Research*, *26*, 597–616.

Pennebaker, J. W. (1996). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*, 162–166.

Price, D. D., Finniss, D. G., & Benedetti, F. (2008). A comprehensive review of the placebo effect: Recent advances and current thought. *Annual Review of Psychology*, *59*, 565–590.

Price, D. D., Milling, L. S., Kirsch, I., Duff, A., Montgomery, G. H., & Nicholls, S. S. (1999). An analysis of factors that contribute to the magnitude of placebo analgesia in an experimental paradigm. *Pain*, *83*, 147–156.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., . . . Engle, R. W. (2012). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*, 359–379.

Rosenthal, D., & Frank, J. D. (1956). Psychotherapy and the placebo effect. *Psychological Bulletin*, *53*, 294–302.

Rudebeck, S. R., Bor, D., Ormond, A., O'Reilly, J. X., & Lee, A. C. H. (2012). A potential spatial working memory training task to improve both episodic memory and fluid intelligence. *PLoS ONE*, *7*, e50431. doi:10.1371/journal.pone.0050431

Schubert, T., & Strobach, T. (2012). Video game experience and optimized executive control skills—On false positives and false negatives: Reply to Boot & Simons (2012). *Acta Psychologica*, *141*, 278–280.

Serfaty, M., Csipke, E., Haworth, D., Murad, S., & King, M. (2011). A talking control for use in evaluating the effectiveness of cognitive-behavioral therapy. *Behaviour Research and Therapy*, *49*, 433–440.

Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, *16*, 97–115.

Stockhorst, U., Steingrüber, H. J., & Scherbaum, W. A. (2000). Classically conditioned responses following repeated insulin and glucose administration in humans. *Behavioural Brain Research*, *110*, 143–159.

Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica*, *140*, 13–24. doi:10.1016/j.actpsy.2012.02.001

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., . . . Cohen, J. D. (2004). Placebo-induced changes in FMRI in the anticipation and experience of pain. *Science*, *303*, 1162–1167.