Commentary

Cogmed training: Let’s be realistic about intervention research

Susan E. Gathercole\(^a,\)*, Darren L. Dunning\(^b\), Joni Holmes\(^c\)

\(^a\) MRC Cognition and Brain Sciences Unit, Cambridge
\(^b\) University of East Anglia
\(^c\) MRC Cognition and Brain Sciences Unit, Cambridge

1. Commentary

There has been an explosion of interest in methods and evidence relating to intensive cognitive training over the past decade. Key abilities for everyday cognitive functioning such as attention, inhibitory control, working memory and processing speed can be enhanced with sustained daily practice (see Rabipour & Raz, 2012, for a comprehensive review). It is therefore unsurprising that the potential to enhance cognitive function has attracted interest both from companies specialising in educational products such as Pearson and from professionals in health and education looking for effective ways to improve outcomes for their pupils and patients.

These potential end-users of cognitive training must look to cognitive scientists both to generate a sound evidence base and to provide a balanced appraisal of it. For those readers who are not familiar with several other recent reviews of working memory training (e.g. Melby-Lervag & Hulme, 2012; Shipstead, Redick, & Engle, 2010, Shipstead, Redick, & Engle, 2012), Shipstead, Hicks and Engle’s article does a valuable job in laying out the full range of published research on the Cogmed training programme in particular. We do, however, have concerns about how realistic are the criteria by which they evaluate (and in most cases, reject) individual studies, and argue here for a broader analysis that weighs up the evidence across the full range of relevant data.

The problem is that assessing the impacts of interventions poses exceptional challenges. A gold standard cognitive intervention would minimally i) employ outcome measures with high construct validity, with multiple measures of each construct, ii) test the degree of generalization iii) randomly allocate participants to condition; iv) include both a no treatment baseline condition and active comparison interventions that control for placebo effects and all potential treatment confounds; vi) have high levels of statistical power.

Statistically significant findings from such a study would engender high confidence for the efficacy of the intervention. There is, however, one large problem. Studies employing these designs are time-consuming and expensive and, until the existing evidence indicates that the intervention passes less stringent empirical tests, the investment is risky. The pragmatics of how best to proceed in developing new treatments are formalized in the clinical trials framework http://clinicaltrials.gov/ct2/info/understand/, and reflected also in the Cooke (2006) pipeline characterizing the progression from basic science through to translation. In the process of developing new treatments, exploration commences with proof of concept studies, progresses through to larger studies designed to calibrate and optimise the parameters of the intervention and test for confounds, and finally, if successful to this point, feeds through to large-scale randomized controlled studies. The probability of a genuine treatment effect increases with each positive outcome, and this justifies the additional investment required for the next phase.

Although studies that do not meet all of the stringent criteria for intervention designs inevitably have relatively elevated rates of false positives, any null results that they generate can have considerable value in allowing candidate interventions to be
dismissed with a relatively high level of confidence. A case history of our own research in this area may help illustrate this. We started off convinced, for what seemed at the time good reasons relating to a large body of prior experimental evidence, that intensive working memory training would not lead to generalized gains in memory performance. Nonetheless, the published results with Cogmed training looked interesting (e.g., Klingberg et al., 2005). The programme also had merits in employing training methods that shared many features of the complex memory span paradigms that have high working memory construct validity (Alloway, Gathercole, & Pickering, 2006; Kane et al., 2004). We therefore decided to see whether the programme worked in the populations of children with working memory problems that interested us.

We didn’t have the necessary resources to set up a large-scale study, but were able to incorporate Cogmed training with a small group of children with ADHD participating in a larger funded project. The design was far more than ideal; it consisted only of a single group that received the intervention and lacked other comparison conditions controlling for factors such as the motivational boost of participating in the training and receiving positive feedback. While positive effects are difficult to interpret with such a simple design, a null effect is not: if performance on other working memory measures is not enhanced following this intervention despite the myriad of potentially beneficial consequences (many of which are not specific to the memory-demanding aspect of the training activities), the probability that it is capable of genuinely and specifically boosting underlying working memory capacity is low. In fact, the outcome was positive (Holmes et al., 2010). The children showed substantial gains in their performance on untrained measures of both short-term and working memory, but not in IQ.

Our first thought was that the training-related increases in memory performance may simply have reflected the beneficial effects for children with deficits in sustained attention of regular engagement in focussed activities in small groups. A further study was conducted with children with low working memory but no diagnosed attentional problems to see if they showed the same response to training. They did: substantial gains were observed in untrained tests of working memory that largely persisted 6 months later. By this time the lack of a control intervention group was a real concern, so we recruited a further group of low-memory children to complete the non-adaptive training programme produced by Cogmed for the purposes of research. The task environment and amount of practice in the tasks is identical to the standard version, but difficulty is below the level of average span and does not increase with success (Klingberg et al., 2005). Our low working memory group showed no training benefits. Thus, the critical feature of the training gains we had observed in the first group of low-memory children appears to have been the extent to which it taxed working memory and drove self-improvement, rather than task practice (Holmes, Gathercole, & Dunning, 2009).

Note that although the same recruitment criteria were applied to the participants in both intervention conditions in the Holmes et al. (2009) study, schools were not randomly allocated to condition, introducing possible sampling differences (Torgerson & Torgerson, 2008). Given the consistency of the enhancements in working memory performance that we had seen in our previous studies, a larger-scale investment in a full double-blind randomized controlled trial (RCT) of Cogmed training in children with low working memory was justified. The Leverhulme Trust funded this next crucial step and, after a gruelling three years, this study is now complete. We have found generalized working memory enhancements in low-memory children (Dunning, Holmes, & Gathercole, undergoing revision), upholding our previous findings using less stringent designs. Without the preliminary evidence that the intervention could work, the full RCT study would not have been justified. However, the predictive value of such data is lost if every study that fails to meet the gold standard criteria is relentlessly rejected.

As Shipstead et al. discuss, many key questions regarding both the fundamental mechanisms underpinning working memory training and its practical applications remain unanswered. We would like to underscore their comments on the need to establish whether the training methods that have been developed really do have the potential to deliver educationally significant gains in academic progress. There have been isolated findings of significant boosts in mathematics (Holmes et al., 2009) and reading comprehension (Dahlin, 2011). However, no replicated pattern of transfer has been established to date, and the reasons for this need to be considered carefully. One possibility is that the training gains in working memory performance that have been observed do not reflect quantitative improvements in the efficiency of the neural substrate as claimed by some (e.g., Westerberg & Klingberg, 2007). Instead, they may arise from the development of relatively task-specific strategies that trainees then fail to apply spontaneously to everyday activities that tax working memory. If this is the case, a fruitful approach may be to provide the trainee with practice in transferring their newly-learned strategies to other situations that more directly simulate the everyday classroom demands on working memory. Optimally exercising potential that has been enhanced by training in the developing brain in this way is considered by many to be an essential step (e.g., Jolles & Crane, 2012). We are therefore working on developing a complex task environment designed to bridge the gap between highly structured and relatively artificial activities that load working memory and its flexible use in educational settings such as, for example, following instructions.

In summary, we argue that in evaluating research on cognitive interventions, it is important to weigh up all of the available data. Cumulative evidence, gauged appropriately, can have immense value. A sceptical stance is vital in science, but it is important also to avoid throwing out the baby with the bathwater.

References


