Can intelligence be increased by training on a task of working memory?

The May 13, 2008 issue of *Proceedings of the National Academy of Sciences* featured a cover article that purported to demonstrate increases in fluid intelligence following training on a task of working memory (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). The authors described their own findings as a "landmark result". Their study was the subject of an introductory comment by Robert Sternberg (2008), as well as articles in the mainstream media, including a lengthy column in a recent edition of *The New York Times* (Wang & Aamodt, 2009).

In view of the potential significance of the study and the quantity of attention it has received, the results have been subjected to remarkably little critical analysis. A close examination of the evidence reported by Jaeggi et al. shows that it is not in fact sufficient to support the authors' conclusion of any increase in their subjects' fluid intelligence.

What Jaeggi et al. reported were modest increases in performance on a test of fluid intelligence following several days of training on a task of working memory. The reported increases in performance are not in question here. But the manner in which the test was administered severely undermines the authors' interpretation that their subjects' intelligence itself was increased.

The subjects were divided into four groups, differing in the number of days of training they received on the task of working memory. The group that received the least training (8 days) was tested on Raven's Advanced Progressive Matrices (Raven, 1990), a widely used and well-established test of fluid intelligence. This group, however, demonstrated negligible improvement between pre- and post-test performance.

The other three groups were not tested using Raven's Matrices, but rather on an alternative test of much more recent origin. The Bochumer Matrices Test (BOMAT) (Hossiep, Turck, & Hasella, 1999) is similar to Raven's in that it consists of visual analogies. In both tests, a series of geometric and other figures is presented in a matrix format and the subject is required to infer a pattern in order to predict the next figure in the series. The authors provide no reason for switching from Raven's to the BOMAT.

The BOMAT differs from Raven's in some important respects, but is similar in one crucial attribute: both tests are progressive in nature, which means that test items are sequentially arranged in order of increasing difficulty. A high score on the test, therefore, is predicated on subjects' ability to solve the more difficult items.

However, this progressive feature of the test was effectively eliminated by the manner in which Jaeggi et al. adminstered it. The BOMAT is a 29-item test which subjects are supposed to be allowed 45 min to complete. Remarkably, however, Jaeggi et al. reduced the allotted time from 45 min to 10. The effect of this restriction was to make it impossible for subjects to proceed to the more difficult items on the test. The large majority of the subjects—regardless of the number of days of training they received—answered less than 14 test items correctly.

By virtue of the manner in which they administered the BOMAT, Jaeggi et al. transformed it from a test of fluid intelligence into a speed test of ability to solve the easier visual analogies.

The time restriction not only made it impossible for subjects to proceed to the more difficult items, it also limited the opportunity to learn about the test—and so improve performance—in the process of taking it. This factor cannot be neglected because test performance does improve with practice, as demonstrated by the control groups in the Jaeggi study, whose improvement from pre- to post-test was about half that of the experimental groups. The same learning process that occurs from one administration of the test to the next may also operate within a given administration of the test—provided subjects are allowed sufficient time to complete it.[1]

Since the whole weight of their conclusion rests upon the validity of their measure of fluid intelligence, one might assume the authors would present a careful defense of the manner in which they administered the BOMAT. Instead they do not even mention that subjects are normally allowed 45 min to complete the test. Nor do they mention that the test has 29 items, of which most of their subjects completed less than half.

The authors' entire rationale for reducing the allotted time to 10 min is confined to a footnote. That footnote reads as follows:

> Although this procedure differs from the standardized procedure, there is evidence that this timed procedure has little influence on relative standing in these tests, in that the correlation of speeded and non-speeded versions is very high ($r = 0.95$; ref. 37).

---

[1] I am grateful to D. Detterman for bringing this issue to my attention.

The reference given in the footnote is to a 1988 study (Frearson & Eysenck, 1986) that is not in fact designed to support the conclusion stated by Jaeggi et al. The 1988 study merely contains a footnote of its own, which refers in turn to unpublished research conducted forty years earlier. That research involved Raven's matrices, not the BOMAT, and entailed a reduction in time of at most 50%, not more than 75%, as in the Jaeggi study.

So instead of offering a reasoned defense of their procedure, Jaeggi et al. provide merely a footnote which refers in turn to a footnote in another study. The second footnote describes unpublished results, evidently recalled by memory over a span of 40 years, involving a different test and a much less severe reduction in time.

In this context it bears repeating that the group that was tested on Raven's matrices (with presumably the same time restriction) showed virtually no improvement in test performance, in spite of eight days' training on working memory. Performance gains only appeared for the groups administered the BOMAT. But the BOMAT differs in one important respect from Raven's. Raven's matrices are presented in a 3×3 format, whereas the BOMAT consists of a 5×3 matrix configuration.

With 15 visual figures to keep track of in each test item instead of 9, the BOMAT puts added emphasis on subjects' ability to hold details of the figures in working memory, especially under the condition of a severe time constraint. Therefore it is not surprising that extensive training on a task of working memory would facilitate performance on the early and easiest BOMAT test items—those that present less of a challenge to fluid intelligence.

This interpretation acquires added plausibility from the nature of one of the two working-memory tasks administered to the experimental groups. The authors maintain that those tasks were "entirely different" from the test of fluid intelligence. One of the tasks merits that description: it was a sequence of letters presented auditorily through headphones.

But the other working-memory task involved recall of the location of a small square in one of several positions in a visual matrix pattern. It represents in simplified form precisely the kind of detail required to solve visual analogies. Rather than being "entirely different" from the test items on the BOMAT, this task seems well-designed to facilitate performance on that test.

More generally, the foregoing considerations suggest a deeper problem with the conclusions presented by Jaeggi et al.: To what extent does improvement on any test of fluid intelligence reflect an increase in actual intelligence rather than merely an increase in test-taking skills? A full analysis of this issue is beyond the scope of the present review, but the methodological challenges involved are formidable and deserve further discussion.

Whatever the meaning of the modest gains in performance on the BOMAT, the evidence produced by Jaeggi et al. does not support the conclusion of an increase in their subjects' intelligence. Their research may be sufficient to encourage further investigation, but any larger inferences are unwarranted.

References

Frearson, W., & Eysenck, H. J. (1986). Intelligence, reaction time (RT), and a new odd-man-out RT paradigm. *Personality and Individual Differences*, 7, 807−818.

Hossiep, R., Turck, D., & Hasella, M. (1999). *Bochumer Matrizentest: BOMAT-Advanced-Short Version.* Göttingen: Hogrefe.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008, May 13). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), 6829−6833.

Raven, J. C. (1990). *Advanced progressive matrices: Sets I, II.* Oxford: Oxford Univ Press.

Sternberg, R. J. (2008, May 13). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), 6791−6792.

Wang, S., & Aamodt, S. (2009, March 10). Guest column: Can we increase our intelligence? *New York Times.* http://judson.blogs.nytimes.com/2009/03/10/guest-column-can-we-increase-our-intelligence/

David E. Moody
Mind Over Math, 530 W. Ojai Ave. #205, Ojai, CA 93023,
United States
E-mail address: Davidem@west.net.

12 April 2009