



Curriculum, Evaluation and Management Centre

PUBLICATION NO.16

EMPOWER AND MONITOR:
THE EM ALGORITHM FOR THE CREATION
OF EFFECTIVE SCHOOLS

C.T. FITZ-GIBBON

1990

Empower and Monitor: The EM Algorithm for the Creation of Effective Schools

CAROL TAYLOR FITZ-GIBBON

Personal experience has an impact and an apparent validity that is not attained by mere information. Perhaps it is for this reason that case studies, anecdotal accounts, or reports of "critical incidents" — all reports, indeed, that give us a sense of "being there" — are appealing and influential. Stake (1978) certainly argued this point.

This paper starts with anecdotes in the hope of providing a feel for one of the topics to be considered: peer tutoring. The paper then moves on to a more scientific mode of argument, the presentation of data on peer tutoring. Finally the need for a kind of monitoring that is professionalizing and empowering is considered.

Thus two rather disparate specific strategies for creating effective schools — peer tutoring and performance monitoring — are seen as aspects of two meta-strategies that are termed "empower" and "monitor".

Anecdotes

Michael was a student in an inner-city school in Los Angeles county (Compton district; names have been changed). He was in the ninth grade and was participating in a peer tutoring project, tutoring fourth-grade students (9-year-olds). A scarf denoting gang membership was hanging out of his back pocket. He had just recently been given the final warning from the school principal, the final warning that precedes expulsion, and he came from one of the classes in the lower half of the school. Michael had just calmed down a 9-year-old who had been upset by Cheryl, her 14-year-old

tutor. The 9-year-old had run away from the tutor, had been pursued by tutor and teacher and been caught by her coat. She was crying and dishevelled.

Researcher: Cheryl's little student, Tracy Robinson, had run away from her and from her teacher, and when I found you, you were talking to her and little Tracy was crying. Would you tell me what was happening and what happened next?

Michael: I just told the little girl, I said why was she crying and Cheryl said that she was running from her and she just didn't want to come to class. So then I told Cheryl, you know, go on back to class and I tried to talk to her. So I told her to put back on her clothes and everything, because little ladies don't want to be looking messed up and all. Then she said that Cheryl was bothering her, so I told her well why don't you come to class with me? At first she didn't want to come. She started to run away from me, so I told her, I say, yeah, well, class is where it's happening right now. And I say Math is just a little humbug. In a few days you get to know. All I want to do is teach you a few things. Then she kept on crying so I wiped away her tears and told her, yeah, I'm going to teach you. Me and you can get together and we could see what we could do about Cheryl and everything — okay? Then I say yes, alright, we can come on off in this little classroom and I see what I could teach you. She said was Cheryl going to be in there? I say yes, she was going to be in, and she says she wasn't coming. So I told her, I say, Well little ladies don't worry about little things like that. Whatever she say she do that's the thing she do. So then she said that she was going to come with me. So we started walking towards the class and I thought everything was alright. I get her teacher and everything and let her know what was happening, and 'fore we knew it we was in there sitting down getting it together.

Note how, despite his complete lack of training in counselling or psychology, Michael had

- * Built up the 9-year-old's sense of personal worth: He addressed her as "little lady", thus complimenting her, offering her a new self-image;
- * Encouraged a sense of personal efficacy: "... Whatever she say she do that's the thing she do";
- * Lowered her anxiety level: "Math is just a little humbug . . .";
- * Offered support: "Me and you can get together . . .";
- * Helped her confront reality: "Yes, she (Cheryl) was going to be in [there]";

What is to be learnt about school effectiveness from this incident?

First, in this particular situation, the student was more effective than the teacher in reaching an upset youngster. Not only can students take on some work from teachers, they may actually be better in some circumstances — and students are available in great abundance. Furthermore, in asking students to help we *empower* them and build up their recognition of the contribution they can make in a school. This empowering is real; it does not have the artificiality of platitudes and consultative councils. Students can take on real tasks to the considerable benefit of all those involved. Opportunities to so empower students should, it can be argued, be *sought out*, not left to chance.

In the anecdote just cited, there was some evidence for a major unanticipated benefit: a change in how Michael was viewed by other students. A student who had observed the incident and watched Michael tutoring fractions to little Tracy remarked "I'd always thought Michael was nothing but a cut-up". Similar changes in perceptions have been noted from teachers observing tutors. Is this, then, the recasting of student roles, a way to change teacher expectations?

Thus major benefits accrue to students who give help: they see their own usefulness, others see them in a better light, and it is to be expected that the tutors learn by having to teach. At least, they should. And the tutees should learn also. That is the expectation. Is it the reality? This is where the second strategy comes in: monitoring. The learning need not be left to chance; it should be *monitored*.

Which leads us to a second anecdote. In the middle of the three-week period during which the 14-year-olds were tutoring 9-year-olds in fractions, the researcher gave the tutees a test on the work so far covered. She circled mistakes and wrote helpful comments and then handed the papers to the tutors when they arrived for the next session a few minutes before the tutees were due. She then turned to put work on the board and shortly thereafter became slightly uneasy. Instead of the usual buzz of conversation there was silence — usually a bad sign. Looking round she found nothing untoward brewing; instead, an intense concentration on the returned papers. If you have not taught 14-year-olds in the bottom half of an eight-stream inner-city junior high school, you may think nothing of such an observation. The researcher had such teaching experience, and she was deeply surprised and pleased. She had many times returned papers only to see them barely glanced at. Fractions, after all, were way down on the students' lists of

priorities in life. The difference in the situation, the factor that led them to study the papers, was that they cared about the learning of the younger pupils. "More of us should come and help. The teachers haven't got the time to help like we can." "If someone had helped me I might have done better."

In short, as many, many others have found through their own compelling personal experience, peer tutoring often works.

If the reader is wondering how it happened that it was possible to report Michael's words absolutely verbatim in the above exchange, the reason is that the researcher had been giving a questionnaire in the school. *Because so many students, even in the ninth grade, could not read easily*, the researcher had recorded the questions on a tape. So the tape recorder was at hand when the incident occurred.

This too is a telling anecdote. How did it happen that so many ordinary, not handicapped, students had not learned to read? How was the school system, along with so many other inner-city school systems, allowed to be so ineffective in such critical ways? Why was this situation allowed to persist, year after year? Many students who were certainly not learning-disabled failed to learn to read.

Is this not a prime example of a failure to set up adequate monitoring systems? It may be protested that California administered regular "achievement" tests. It did. Standardized, multiple-choice, speeded tests. The students concerned came out in the lowest percentiles, year after year, undifferentiated. A teacher who had actually been successful would never find this out from the kind of monitoring conducted, for two reasons. First the test, although called an "achievement" test, actually measured ability, especially the ability to respond quickly to multiple-choice items. Because of the way the items were developed and chosen, it was an ability measure rather than a measure of the hard-won skills of the classroom, the teachable skills, the "achievement" a teacher can effect. Second, the monitoring fed back only school and county averages, so that any effective teaching that was occurring was invisible, swamped by the general pattern.

Data

As just described, participant observation by an experienced teacher had suggested that peer tutoring changed the behavior of tutors, producing helping behaviors and time-on-task. It appeared more successful than any

other strategy the teacher had tried. Are there any data to support such a view?

The research from which the above anecdotes were drawn was actually conducted as four field experiments with randomized control groups. The effect sizes (ESs) are shown in figure 1.

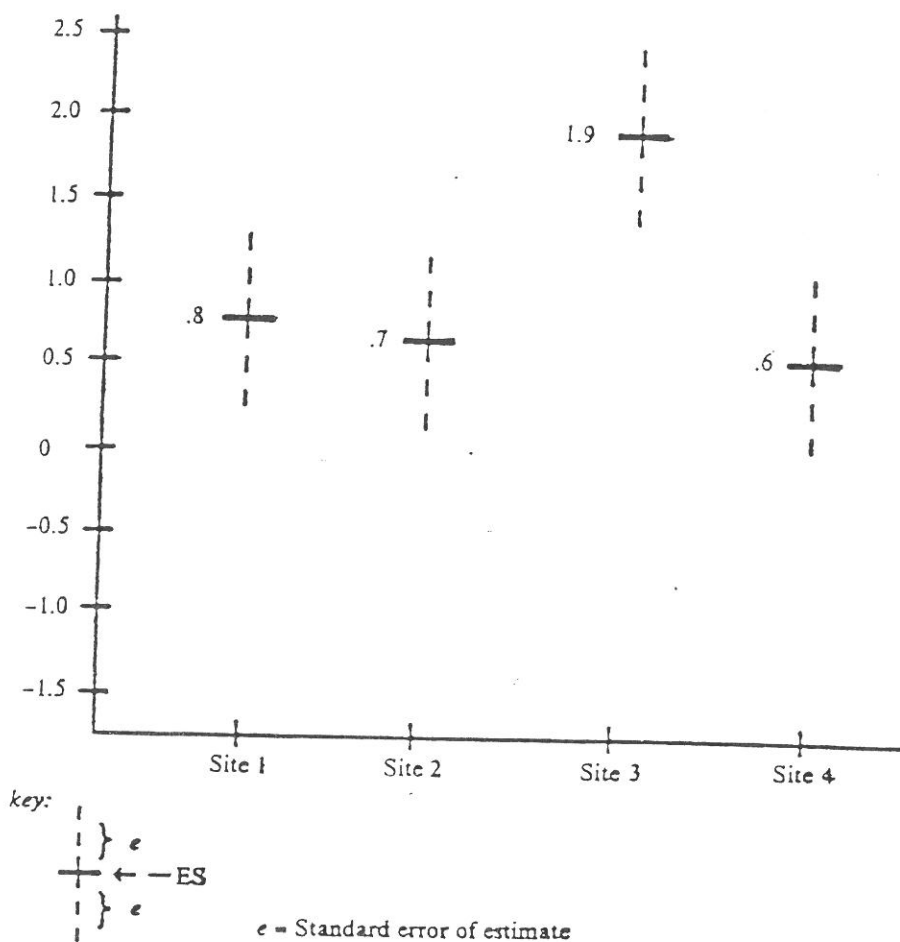


Figure 1. Effect sizes from four controlled experiments in peer tutoring.

When compared with "normal instruction" the ES was about .7 on average in this mathematics peer tutoring project. (The exceptionally large

ES in site 3 was probably due to the fact that the teacher of this class did not teach fractions during the three weeks when the tutors were removed to tutor fractions, as was the case in the other three classes. She said they had already done fractions. The class therefore constituted more of a no-treatment control group than a competing-treatment control group.)

Those field trials were reported in 1975. Evidence as to exactly how effective peer tutoring is has continued to surface and to give remarkably consistent estimates for mathematics:

- * In 1977, Hartley summarized the results of four strategies for improving the effectiveness of mathematics instruction: computer-assisted learning, mastery learning, programmed learning and cross-age tutoring. Cross-age tutoring "won hands down" (Glass 1978) with an average ES of .60. Cross-age tutoring was both the most effective and the most cost-effective.
- * By 1982, Cohen et al. had been able to find 64 controlled experiments in cross-age tutoring. The average effect size for mathematics was .62 for tutors and .60 for tutees.
- * In 1984, Levin et al. reported yet another meta-analysis, which included peer tutoring as one of the treatments, comparing it with increased instructional time and reduced class sizes and computer-assisted instruction. Once again cross-age tutoring "won hands down". Effect sizes for mathematics were reported as 1.02 and .91 in grades 2 and 3 for tutees receiving tutoring from upper grade students.

These consistently positive results from well-controlled studies are the crux of any argument in favour of "empowering" students in peer tutoring projects. Many interventions have their fervent supporters and abundant amounts of rhetoric in their favour. Few have the consistent support of such varied instances of field experiments from many different sources.

However, no matter how strong the research evidence, the generalizability of the results usually remains to some extent a matter for conjecture. Will the methods work in your own particular situation: with your particular staff, given your type of resources and your particular students? The only answer is to try it and see. This is where the existence of a monitoring system becomes a highly valuable asset. It provides, or should provide, the chance for what Cooley, in a far-sighted article, called "Monitoring and Tailoring" (Cooley 1983). For monitoring to be useful in this way it must have certain characteristics.

Desiderata in a Monitoring System

1. *A monitoring system should feed back information to a managed unit.* If there is no chance of change on the basis of feedback from the monitoring, then the monitoring cannot improve the delivery of education. Change requires that someone who is in receipt of the findings of the monitoring should be able to make changes or cause changes to be made. Since there are many levels of management there will need to be many monitoring systems, each dealing with a set of inputs and outputs that are likely to be related to a set of processes.

2. *A monitoring system should have an explicit rationale.* The rationale needed is essentially a model of the major variables in a monitored system. The outcomes must be specified, and the inputs that affect these outcomes must be measured in order to make fair comparisons. Process variables should be assessed in order to inform efforts for improvement. In other words, the concept of improvement from monitoring requires that the monitoring be based on a *model* in which there are potentially *causal* relationships that are tapped by the monitoring.

3. *The system should be a negotiated one.* The monitoring system should have been negotiated and accepted by students, teachers, and administrators. At the very least the dependent variables should be outcomes that reflect the agreed purposes of the unit of the educational system being monitored.

4. *A monitoring system should have positive behavioral implications.* Measurement is intrusive, reactive. The very fact of measuring certain outcomes draws attention to these outcomes. This situation can be used to advantage. (For example, a beginning has been made in measuring incidents of students feeling unsafe or reporting racial insults. The very fact that these aspects of school experience are being assessed makes it likely that more attention will be directed towards dealing with such problems of school life.)

5. *A monitoring system should not unduly upset the system it is monitoring.* If the collection of the data takes up inordinate amounts of time, usurps educationally important activities, or stresses participants, then we need to look again at the monitoring. (The amazing plan for dozens of Standard Assessment Tasks (SATs) to be administered by teachers in every subject for the newly introduced National Curriculum in the U.K. is probably a case in point. The inordinate demands and the threat of publication of the results classroom by classroom will almost certainly cause counterproductive stress

in the system. It is predictable that SATs will pass away — though painfully, like kidney stones.)

Is Monitoring Empowering?

Too often monitoring carries the implication of mechanistic, judgmental systems. The thoughtless imposition of systems of surveillance could be far from empowering (except to the extent that teachers might verify again their power to disrupt or subvert systems of which they do not approve). There are plenty of warnings in the social psychology literature. For example, some of the literature, in particular attribution theory, generates questions concerning intrinsic motivation and the effects of surveillance (e.g., Butler 1988). On the other hand there is evidence that feedback systems can be motivating (Parsons 1974). It may be that the impact of monitoring systems varies with the kind of feedback and the prevailing perceived ethos in the system.

Experience with a project monitoring A-levels encourages the belief that when process data are emphasized and stress is laid on the generation of ideas for improvement, there is considerable interest and no objections to the collection and reporting of such data. Since 1983 a project at Newcastle University (U.K.) has been collecting data and feeding them back to school departments in the form of "league tables" (Fitz-Gibbon 1985, 1990a, 1990b). No school or college has withdrawn from the system, which now includes almost every institution offering A-levels in northeastern England. The project is supported by seven local education authorities.

Given teachers who receive feedback from a monitoring system that is informative, fair, and agreed, and given that they have some input into the design of the monitoring for the next cycle, there is hope that monitoring can prove empowering. Teachers will have access to as good a set of data as anyone. They will have a means to experiment with teaching methods to see if the changes they make appear to influence outcomes. They will know when interventions are needed to maintain quality and can thus operate a self-monitoring profession. The ideal put forward by Campbell in "Reforms as Experiments" (1969) might be realized. In such a situation one of the instructional processes found to be effective will very likely be peer tutoring.

References

- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58, pp. 1-14.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24, pp. 409-429.
- Cohen, P. A., Kulik, J. A. and Kulik, C. L. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 19(2), pp. 237-248.
- Cooley, W. W. (1983). Improving the performance of an educational system. *Educational Researcher*, 12(6), pp. 4-12.
- Fitz-Gibbon, C. T. (1985). A-level results in comprehensive schools: The Combse project, year 1. *Oxford Review of Education*, 11(1), pp. 43-58.
- Fitz-Gibbon, C. T. (1990a). An Up-and-Running Indicator System. In Fitz-Gibbon, C. T. (ed.), *Performance Indicators: A BERA Dialogue*. Clevedon, Avon: Multi-lingual Matters.
- Fitz-Gibbon, C. T. (1990b). Performance indicators: Educational considerations. In Levacic, R. (ed.) *Financial Management in Education*. Milton Keynes: Open University Press.
- Fitz-Gibbon, C. T., Tymms, P. B., and Hazelwood, R. D. (1989). Performance indicators and information systems. In Reynolds, D., Creemers, B. P. M., and Peters, T. (eds.), *School Effectiveness and Improvement*. Cardiff and Groningen: University of Wales College of Cardiff and RION Institute for Educational Research.
- Glass, G. V. (1978). Reply to Mansfield and Busse. *Educational Researcher*, 7(1), p. 3.
- Hartley, S. S. (1977). A meta-analysis of effects of individually paced instruction in mathematics. Doctoral dissertation, University of Colorado.
- Levin, H. M., Glass, G. V., and Meister, G. R. (1984). *Cost-effectiveness of Four Educational Interventions*. Project Report No. 84-A11. Stanford: Institute for Research on Educational Finance and Governance, Stanford University.
- Levin, H. M., Glass, G. V., and Meister, G. R. (1987). Cost effectiveness of computer-assisted instruction. *Evaluation Review*, 11(1), pp. 50-72.

- Parsons, H. M. (1974). What happened at Hawthorne? *Science*, 183, pp. 922-932.
- Stake, R. L. (1978). The case study method in social enquiry. *Educational Researcher*, pp. 5-8.