Curriculum, Evaluation and Management Centre

PUBLICATION NO.21

PEER AND CROSS AGE TUTORING

C.T. FITZ-GIBBON

Encyclopedia of Educational Research
1992

Shumaker, S. A., & Brownell, A. (1984). Towards a theory of social support: Closing conceptual gaps. *Journal of Social Issues, 40*, 11–36.

Varenhorst, B. B. (1974). Training adolescents as peer counselors. *Personnel and Guidance Journal, 53*, 271–275.

Varenhorst, B. B. (1984). Peer counseling: Past promises, current status and future directions. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (pp. 716–750). New York: John Wiley & Sons.

PEER AND CROSS-AGE TUTORING

The term *peer tutoring* (PT) has been applied to various activities. The common threads are the use of *peers*, that is, others of similar status and situation, and *tutoring*, which implies individualized attention to a learner, the *tutee*.

Individual tuition, a class size of one, has always been prized. Members of the aristocracy were usually educated by personal tutors, and even today, if a student wishes to excel or catch up, a private tutor is often seen as the answer. If this widely held belief in the effectiveness of one-to-one tutoring were correct and schools wished to provide every student with a personal tutor, on occasion, the only feasible approach in terms of costs would be to use students as the tutors. But are students effective tutors? There is an increasing body of evidence that they are and that, moreover, the tutors themselves may benefit from having to tutor. The claims for positive outcomes extend beyond the cognitive and into other domains, such as attitudes toward self and others.

The evidence for cognitive and noncognitive benefits is reviewed separately. These reviews are followed by consideration of some important issues emerging from the research literature.

Cognitive Benefits

The learning outcomes of PT have been widely researched, primarily in the basic-skill areas of reading and mathematics (Allen, 1976; Goodlad & Hirst, 1989, 1990; Topping, 1988). Applications of PT in curriculum areas other than basic skills are increasing, as, for example, in science (Tobias, 1986) and in language learning (Walz, 1986).

A feature of PT research is the extensive use of experimental designs in school settings and the variety of sources from which data are available. This feature has made possible a number of meta-analyses (see Table 1). Hartley (1977) located 73 studies of tutoring being used to deliver individualized mathematics instruction. Cohen, Kulik, and Kulik (1982) located 65 independent evaluations of school tutoring programs and Cook, Scruggs, Mastropieri, and Casto (1986) computed 74 effect sizes from 19 articles evaluating the impact of having handicapped students serve as tutors. In a study of the cost effectiveness of four interventions, Levin, Glass, and Meister (1987) chose a large-scale, established tutoring program at Boise, Idaho (U.S.) as representative of tutoring projects. The results presented in Table 1 show positive effect sizes, thus indicating the general effectiveness of tutoring projects. Effect sizes for mathematics were substantial and were about twice as large as those for reading.

The effect sizes presented in Table 1 should be compared to those from other forms of intervention designed to improve basic skills: 0.23 for computer-assisted instruction (Hartley, 1977) and only 0.12 on average for the computer-assisted instruction examined by Levin,

TABLE 1. *Results of literature surveys of the cognitive outcomes[a] of tutoring*

Survey	*Tutors' outcomes*		*Tutees' outcomes*		*Type of tutor*
	Reading	*Mathematics*	*Reading*	*Mathematics*	
Hartley (1977)	—[b]	0.58	—	0.54	Adults
				0.52	Same age
				0.79	Cross age
Cohen, Kulik, & Kulik (1982)	0.21	0.62	0.29	0.60	[c]
Levin, Glass, & Meister (1987)	—	—	0.48	0.97	Cross age
	—	—	0.38	0.67	Adults[d]
Cook, Scruggs, Mastropieri, & Casto (1986)	0.30	0.67	0.49	0.85	Cross age

[a]Outcomes are expressed as effect sizes or standardized mean differences between the tutoring group and control group.

[b]Hartley's study was concerned only with individualized mathematics instruction.

[c]Combining across reading and mathematics, the results for tutors were 0.28 for same age and 0.35 for cross-age situations. For tutees the corresponding figures were 0.29 and 0.49.

[d]Cross-age tutoring was used for grades 2 and 3 and adult tutoring for grades 4, 5, and 6.

Glass, and Meister (although the findings have been questioned by Niemiec, Blackwell, & Walberg, 1986; Levin, Glass, & Meister, 1986). Other comparative findings were 0.16 for mastery learning, 0.22 for reducing class size from 35 to 20, and 0.03 for increasing instructional time.

As Table 1 shows, differences between effect sizes appeared to be related to implementation characteristics. For example, adult tutors (usually paid paraprofessionals) tended to be less effective than pupil tutors. Another replicated difference was between same-age and cross-age tutoring. The average effect sizes appeared to favor cross-age projects. These results, however, could have been chance relationships. Attributions are difficult, because these implementation characteristics were not manipulated variables. Nevertheless, the patterns suggest that careful consideration should be given to these implementation characteristics in designing tutoring projects.

A major threat to the interpretation of these generally strong, positive findings arises from consideration of the nature of the control groups' activities, including the time allocated and the materials used. Experiments comparing PT directly with a variety of feasible alternative teaching strategies are still needed (cf. Scruggs & Richter, 1988).

Interpreting outcomes of PT projects involves an important distinction between *Tutorial-service* and *Learning-by-tutoring* projects. In the latter case, tutors teach a topic that they themselves need to learn and practice; tutors and tutees alike are expected to show learning gains. However, when tutors teach a topic that they already know thoroughly, they are essentially providing a service, and such projects might be designated tutorial-service projects; the tutors may be volunteers or paid. Two major tutorial-service projects have been in operation for several years. In London, hundreds of university students have assisted in school classrooms (Goodlad & Hirst, 1989), and in Israel 12,000 university and college students per year have served as personal tutors to disadvantaged school students (Davis, Snapiri, & Golan, 1984; Fresko & Eisenberg, 1985).

Noncognitive Benefits

The effects of PT on attitudes and behaviors, i.e., the noncognitive benefits of PT, are sometimes considered more important than the cognitive benefits. Numerous applications of the technique have been aimed at improvements in areas other than learning outcomes.

Attitudes Toward Self and Others. Bringing students together to work on tutoring may improve interpersonal relationships. This result can be particularly valuable in situations where there are barriers and/or hostilities. That is, interethnic friendships may be fostered (Fitz-Gibbon, 1983; Slavin, 1979) and other barriers of gender, socioeconomic status, or handicap may be bridged. Applications in which special-education students have served as tutors and/or tutees have been reviewed by Osguthorpe and Scruggs (1986) and Eiserman, Shisler and Osguthorpe (1987). Belief in the beneficial effects of tutoring on self-esteem, self-concept, and sociometric ratings is widespread but has limited, though by no means zero, empirical support. Particularly difficult is the question of the generalization of attitudes from the tutoring situation, and from the particular students involved, to other situations and other students.

To ask students to tutor is to imply that the students have something to offer tutees. The request itself denotes confidence in the tutors. If the tutors then find that they can indeed help effectively, the experience should enhance their self-esteem. Obtaining firm evidence for improved self-esteem (Cohen, Kulik, & Kulik, 1982) has proved difficult, although there are some indications of positive outcomes (e.g., Maheady & Sainato, 1985).

Behaviors. Schunk (1987) reviewed the use of peers as models to promote behavioral change. The use of behavioral techniques such as reinforcement has been particularly notable in special-education settings (Strain, 1981). Greenwood, Carta, and Hall (1988) concluded that peer-mediated procedures for classroom management could be effective, noting particularly the opportunities they provide for "academic and social responding, cooperation, and peer conflict resolution" (p. 271).

Health has been a major concern for a number of peer-mediated interventions. One such intervention started in 1979 is Child-to-Child, an internationally supported project to train children to promote health-supporting practices in developing countries. As part of the project, materials on oral rehydration, care of sick children, and the like have been produced. (Aarons, Hawes, & Gayton, 1979).

Drug-taking behavior, including smoking, has been the target of some peer counseling projects, which were meta-analyzed by Bangert-Drowns (1988). As in the meta-analyses referred to in Table 1, peers were more effective than adults, with effect sizes for attitudinal changes being 0.64 for peers and 0.26 for adults. For drug-taking behaviors, typically very refractory, effect sizes were 0.31 for peer-taught students and virtually zero for adult-taught students.

Attitudes to School. By introducing an enjoyable activity, providing good role models, and giving emotional support, those involved in a tutoring project may improve bonding to the school and reduce the sense of deindividuation or alienation that is thought to precede much antischool and antisocial behavior (Hawkins, Doueck, & Lishner, 1988). Indeed Washburn (1975), an anthropologist, argued that problems arise in schools "from the loss of the traditional folk-learning situation in which learning depends on identification, emotion, and clearly visible goals" (p. 8). He suggested that "the point of teaching by peers is to restore the conditions of learn-

ing that are natural to man, the kind of situations in which the human brain evolved" (p. 8).

Social Skills and Social Status. In some projects the concern may be predominantly with the personal and social development of the participants, whether it be overcoming shyness or remediating antisocial or anti-school behaviors. Tutors are sometimes specifically trained in social skills such as addressing the tutee by name, smiling, and praising (Maheady & Sainato, 1985; Scruggs & Richter, 1988; Wheldall & Mettem, 1985).

Impact on Teachers. Comments from tutors frequently indicate that taking on a tutoring role enhances their empathy toward teachers. However, there appears to have been no research into this or any other possible effects on teachers of adopting PT as a teaching strategy. As for recruitment into the teaching profession, PT enables tutors to explore the possibility of teaching as a career, but the impact—positive or negative—appears not to have been assessed.

Issues Arising from the Literature on Peer Tutoring

How does PT produce positive outcomes? What are the relevant theories and mechanisms?

Some educators argue that tutees benefit from individualized attention and close monitoring. Tutees may also understand an older pupil better than they understand a teacher or other adult, because the older pupil's language may be closer to that of the tutee. Additionally, tutees may feel more able to ask questions of a tutor than of a teacher or other adult. Citing Piaget and Vygotsky, some authors have emphasized the importance of social interaction and cognitive conflict in learning (Foot, Shute, Morgan, & Barron, 1990). They have also looked at process rather than outcomes and have taken a more pessimistic view of the likelihood of learning benefits than that generally taken by school personnel.

Considering tutors, enhanced motivation and verbalization, and even simple arousal, are cited as sources of frequently improved learning (cf. Benware & Deci, 1984; Light & Glachan, 1985). There are, however, two situations in which tutors are not expected to show increased learning: if they are teaching work that they already fully understood, as in tutorial-service projects, or if they do not take on the tutoring role, i.e., if they do not try to explain the work to the tutee. Failures are publicized less often than successes but there are one or two warnings available. Vedder (1985) reported a large-scale, controlled, field experiment in collaborative learning (a kind of same-age PT) and concluded that there had been a failure of implementation. He identified problems similar to those cited by Foot, Shute, Morgan, and Barron (1990), such as tutors who were not sensitive to tutees' learning or who seldom gave good explanations. Reviewing this work Slavin (1987) stated that, for effectiveness, collaborative learning needs reward structures that provide the incentive for students "to try to ensure that their groupmates have learned the material" (p. 63).

This need for an incentive may be one of the reasons for the smaller effect sizes, noted earlier, associated with same-age as opposed to cross-age projects. Given a cross-age pairing, with a sufficiently large age difference, the younger child may function as the incentive, without the need for an external and artificial reward structure. Cross-age projects may have a further advantage for tutees by avoiding the dangers associated with receiving help: "aversive feelings of incompetence, insecurity, indebtedness and dependency" (DePaulo et al., 1989, p. 423). Help from an older student is more acceptable than from one of the same age.

Scruggs and Osguthorpe (1986) working with learning disabled and behaviorally disordered students compared cross-age tutoring with same-age tutoring and found similar cognitive outcomes but better attitudinal results in the cross-age experiment. Working with girls, Ludeke and Hartup (1983) demonstrated that younger tutees evoked more supportive teaching strategies from their tutors.

Whether same-age projects are indeed less effective for some reason than cross-age projects remains to be seen. Meanwhile, in considering which kinds of PT to implement, attention must be paid to ease of organization. Same-age projects within a single classroom are more easily organized than cross-age projects involving two or more classrooms. One same-age technique that has received some support is *reciprocal peer tutoring* (RPT), in which the pupils work in pairs and alternate between being the tutor and the tutee (Fantuzzo, Riggio, Connelly, & Dimeff, 1989; Moore, 1988). Another is *classwide peer tutoring* (Greenwood et al., 1987; Delquadri, Greenwood, Wharton, Carta, & Hall, 1986), which is similar to RPT in that students alternate tutor and turee roles, but which also involves a structured set of procedures growing out of behavioral analysis literature. The system includes rewards for appropriate tutoring behaviors, defined times for tutor and tutee roles, constant changing of pairs, and a competitive points system for individuals and pairs.

Conclusions

The cognitive benefits of PT are well documented by field experiments, making it a procedure that can be confidently recommended for use over short periods of time (such as a few weeks). Some theoretical questions have been raised following detailed investigation of the processes involved. Answering these questions requires research into the extent to which the effects are the result of various underlying mechanisms (Cohen, 1986; Goodlad & Hirst, 1989).

Various applications have been directed at changes in behavior and attitudes. Although some findings are less

clear than in the cognitive area, many promising results lend support to plans to use peers as tutors.

Carol Taylor Fitz-Gibbon

See also Class Size; Cooperative Learning; Motivation.

REFERENCES

Aarons, A., Hawes, H., & Gayton, J. (1979). *Child-to-child.* London: Macmillan.

Allen, V. (Ed.). (1976). *Children as teachers: Theory and research on tutoring.* New York: Academic Press.

Bangert-Drowns, R. L. (1988). The effects of school-based substance abuse education—A meta-analysis. *Journal of Drug Education, 18*(3), 243–264.

Benware, C. A., & Deci, E. D. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21*(4), 755–765.

Cohen, J. (1986). Theoretical considerations of peer tutoring. *Psychology in the Schools, 23,* 175–186.

Cohen, P. A., Kulik, J. A., & Kulik, C. L. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*(2), 237–248.

Cook, S. B., Scruggs, T. E., Mastropieri, M. A., & Castro, G. C. (1986). Handicapped students as tutors. *The Journal of Special Education, 19*(4), 483–492.

Davis, D., Snapiri, T., & Golan, P. (1984). *A survey of tutoring activities in Israel and associated evaluation studies.* Jerusalem: The Hebrew University of Jerusalem.

Delquadri, J., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Classwide peer tutoring. *Exceptional Children, 52*(6), 535–542.

DePaulo, B. M., Tang, J., Webb, W., Hoover, C., Marsh, K., & Litowitz, C. (1989). Age differences in reactions to help in a peer tutoring context. *Child Development, 60,* 423–439.

Eiserman, W. D., Shisler, L., & Osguthorpe, R. T. (1987). Handicapped students as tutors: A description and integration of three years of research findings. *B.C. Journal of Special Education, 11*(3), 215–231.

Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: A component analysis. *Journal of Educational Psychology, 81*(2), 173–177.

Fitz-Gibbon, C. T. (1983). Peer-tutoring: A possible method for multi-ethnic education. *New Community, 11*(1/2), 160–166.

Foot, H. C., Shute, R. H., Morgan, M. J., & Barron, A. (1990). Theoretical issues in peer tutoring. In H. C. Foot, M. J. Morgan, & R. H. Shute (Eds.), *Children helping children.* New York: John Wiley & Sons.

Fresko, B., & Eisenberg, T. (1985). The effect of two years of tutoring on mathematics and reading achievement. *Journal of Experimental Education, 53*(4), 193–201.

Goodlad, S., & Hirst, B. (1989). *Peer tutoring: A guide to learning by teaching.* London: Kogan Page & New York: Nichols.

Goodlad, S., & Hirst, B. (Eds.) (1990). *Explorations in peer tutoring.* Oxford: Basil Blackwell.

Greenwood, C. R., Carta, J. J., & Hall, R. V. (1988). The use of peer tutoring strategies in classroom management and educational instruction. *School Psychology Review, 17*(2), 258–275.

Greenwood, C. R., Dinwiddie, G., Bailey, V., Carta, J. J., Dorsey, D., Kohler, F. W., Nelson, C., Rotholz, D., & Schulte, D. (1987). Field replication of classwide peer tutoring. *Journal of Applied Behavior Analysis, 20*(2), 151–160.

Hartley, S. S. (1977). *A meta-analysis of effects of individually paced instruction in mathematics.* Unpublished doctoral dissertation. University of Colorado.

Hawkins, J. D., Doueck, H. J., & Lishner, D. M. (1988). Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. *American Educational Research Journal, 25*(1), 31–50.

Levin, H. M., Glass, G. V., & Meister, G. R. (1986). The political arithmetic of cost-effectiveness analysis. *Phi Delta Kappan, 68*(1), 69–72.

Levin, H. M., Glass, G. V., & Meister, G. R. (1987). Cost effectiveness of computer-assisted instruction. *Evaluation Review, 11*(1), 50–72.

Light, P., & Glachan, M. (1985). Facilitation of individual problem solving through peer interaction. *Educational Psychology, 5*(3,4), 217–225.

Ludeke, R. J., & Hartup, W. W. (1983). Teaching behaviors of 9- and 11-year-old girls in mixed-age and same-age dyads. *Journal of Educational Psychology, 75*(6), 908–914.

Maheady, L., & Sainato, D. M. (1985). The effects of peer tutoring upon the social status and social interaction patterns of high and low status elementary school students. *Education and Treatment of Children, 8*(1), 51–65.

Moore, P. J. (1988). Reciprocal teaching and reading comprehension: A review. *Journal of Research in Reading, 11*(1), 3–14.

Niemiec, R. P., Blackwell, M. C., & Walberg, H. J. (1986). CAI can be doubly effective. *Phi Delta Kappan,* Vol. 67, 750–751.

Osguthorpe, R. T., & Scruggs, T. E. (1986). Special education students as tutors: A review and analysis. *Remedial and Special Education (RASE), 7*(4), 15–25.

Schunk, D. H. (1987). Peer models and children's behavioral change. *American Educational Research Journal, 57*(2), 149–174.

Scruggs, T. E., & Osguthorpe, R. T. (1986). Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. *Psychology in the Schools, 23,* 187–193.

Scruggs, T. E., & Richter, L. (1988). Tutoring learning disabled students: A critical review. *Learning Disability Quarterly, 11*(3), 274–286.

Slavin, R. E. (1979). Effects of biracial learning teams on cross-racial friendships. *Journal of Educational Psychology, 71*(3), 381–387.

Slavin, R. E. (1987). Cooperative learning in geometry: 180 degrees off. [Review of *Cooperative Learning*]. *Contemporary Psychology, 32*(1), 62–63.

Strain, P. S. (Ed.). (1981). *The utilization of classroom peers as behavior change agents.* New York: Plenum.

Tobias, S. (1986). Peer perspectives on the teaching of science. *Change, 18*(2), 36–41.

Topping, K. (1988). *The peer tutoring handbook: Promoting cooperative learning.* London and Sydney: Croom Helm; Cambridge, MA: Brookline Books.

Vedder, P. (1985). *Cooperative learning: A study on processes and effects of cooperation between primary school children.* Groningen, Netherlands: Rijksuniversteit Groningen-Vakgroep Interdisciplinaire Onderwijskunde.

Vedder, P. (1987). Cooperative learning. *Contemporary Psychology*, 32(1), 62–63.

Walz, J. C. (1986). Increasing student-talk time in the foreign language classroom. *Canadian Modern Language Review*, 42(5), 952–967.

Washburn, S. L. (1975). Evolution and learning: A context for evaluation. *National Elementary School Principal*, 54(6), 4–10.

Wheldall, K., & Mettem, P. (1985). Behavioural peer tutoring: Training 16-year-old tutors to employ the "pause, prompt and praise" method with 12-year-old remedial readers. *Educational Psychology*, 5(1), 27–44.

PERFORMANCE MEASUREMENT

Performance measurement calls for examinees to demonstrate their capabilities directly, by creating some product or engaging in some activity. Compared to objective tests, performance measurements are generally more complex and costly to administer and score—and often have lower reliability. A performance test may be based on only one or a few items (exercises), although each may be rated or scored with respect to several performance dimensions. The cost and complexity of performance measurements are usually justified by evidence or arguments that they are more valid than alternatives for specific applications (Priestley, 1982).

In educational settings, performance measurement may also be justified on the basis of its pedagogical value and its probable impact on curriculum and instruction (Wiggins, 1989). Objective tests may be valid indicators of important knowledge and skills, but regular classroom instruction that is aimed directly at answering test questions, and no more, is generally regarded as a distortion of the curriculum. Instruction aimed directly at improving performance on well-designed performance tests may be more closely aligned to important curriculum goals and is likely to be of greater value than practice in answering objective test questions (Stiggins & Bridgeford, 1985).

This article first discusses the definition of performance measurement, then summarizes its major educational applications. In the final section, some considerations in the design and validation of performance measurements are presented.

Definition

The term *performance test* is often reserved for assessments that involve sampling and quantifying some observable behavior or product of behavior valued in itself, outside of the test setting. Writing samples may be used to assess writing ability; an object created in a machine shop may be evaluated to assess skill in using power tools; or a teacher's classroom performance may be observed and rated to assess teaching proficiency. A broader, alternative definition includes any tests in which the stimuli presented or the responses elicited emulate some aspects of nontest settings. For instance, a test in which teacher examinees observe and discuss a videotape of someone presenting a lesson would not be a performance test by the stricter definition, because watching such videotapes is not itself an important part of teaching. It might be considered a performance test by the broader definition, however. Historically, various nonverbal psychological tests have also been referred to as performance tests. Examples include the Porteus Maze Test, or the Block Design subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R).

Perhaps the narrowest definition of performance measurement is the sampling and quantification of some behavior that would occur whether it were being assessed or not. The routine activities of workers on the job might be observed and rated, for example. If such ratings faithfully reflect the aspects of job performance to be assessed, the validity of such measurements could be argued in part on the basis of a direct statistical generalization from the sample of behavior rated to the universe of behavior represented. Although performance measurements based on such direct samples are often impractical, unreliable, or otherwise unsound, they are ideal in the narrow sense that both the stimuli (i.e., the context, instructions, and materials that define what the examinee is to do) and the responses given precisely represent the criterion to be predicted.

Other performance measurements are based on performances or products elicited solely for the purpose of evaluation. In writing assessments, examinees generally respond to assigned prompts, and their writing serves no purpose other than demonstrating their proficiency. Teacher performance measurements have been devised in which each examinee is given the task of teaching some specific content to a small group of students and then is evaluated according to the students' posttest performance, or on the basis of pretest to posttest gain scores (McNeil, Alkin, & Klein, 1974; Popham, 1971).

In tasks designed exclusively for assessment purposes, duplicating all the features of naturally occurring tasks may be unnecessary, as well as difficult or costly. Tasks used for performance measurement are often simpler in many respects than the corresponding tasks in natural settings. Both the stimuli provided and responses requested may be more or less artificial—dimensions that Fitzpatrick and Morrison (1971) referred to as *stimulus fidelity* and *response fidelity*. Thus there may not be any sharp distinction between performance measurements and other forms of measurement. A broader definition of performance measurement simply includes assessments high in stimulus or response fidelity.

Consider, for example, the assessment of a teacher's