What works in Special Education

Presentation to the Special Education Conference,

May, 1992

I thank you for inviting me today as I feel like many parents must when facing a group of Special Education experts. I know something about special education, I recognize that you know more than I do but, like parents, I am going to constantly remind you that you must continue to know more than me or I will confront you with my knowledge. So, when your President asked me to present a session here today I decided to enter the lion's den with a provocative title, which invites the answer: Ritalin, more PD days, and long service leave. This response reminds me that you are constantly in touch with reality, and perhaps this is why your President asked a non-special educator to address you: You wanted some light relief from your usual, as I'm probably not in touch with reality.

Let me commence with my claim today and then defend it.

The effects of schooling that are achieved by special education students matches that of "normal" children both in

a. the degree of effects you attain in achievement

b. the types of procedures which are effective in enhancing achievement.

Let me hasten to say what I am thereby not saying.

I am not saying that, therefore, special education teaching and "normal" teaching can be achieved by the same person, under the same conditions, necessarily in the same classroom. The effects I am comparing have been achieved typically under conditions of exclusivity. That is, in special education programs and not in "normal" classrooms.

To make my claims I want to commence by referring you to an study that I have been refining over the past five years (Hattie, 1992).

Despite the obvious importance of the question -- "Measuring the effects of schooling" -- there have been few attempts to answer it with empirical data. There are many models of schooling (e.g., Bloom, 1976, Carroll, 1963, and see Hattie,1987 for comparisons), much discussion about desirable outcomes, but little evaluation of the models. I recognize that it is difficult to devise an experiment or directly test the effects of schooling.

The beginning of the answer to the question as to the effects of schooling is to ask "What are the 'typical' effects of schooling?" and then to use this typical effect as a benchmark. The problem is how to ascertain "typical effects" given the myriad of effects on schools, different teachers, subjects, school administration systems, ages of students, and other moderators such as gender, prior ability, quality of instruction and teaching styles. The first requirement is a continuum on which the effects of schooling, including the typical effect, can be summarized where 0 means that there is no effect from introducing some teaching package, innovation, or effect on schooling. A negative effect indicates that the innovation has a decreased effect on achievement, and a positive effect indicates that the innovation has as increased effect on achievement. For the present, the model is constrained to achievement outcomes, but the continuum can be generalized to other outcomes of schooling.

The next requirement is to formulate an appropriate scale and it is recommended that the scale is expressed in effect-sizes. An effect-size provides a common expression of the magnitude of study outcomes for all types of outcome variables, such as school achievement. An effect-size of 1.0 indicates an increase of one standard deviation, typically associated with advancing children's achievement by one year, improving the rate of learning by 50%, or a correlation between some variable (e.g., amount of homework) and achievement of approximately .50. When implementing a new program, an effect-size of 1.0 would mean that approximately 95% of outcomes positively enhance achievement, or average students receiving that treatment would exceed 84% of students not receiving that treatment. Cohen (1977) argued that an effect-size of 1.0 would be regarded as large, blatantly obvious, grossly perceptible, and he provided examples such as the difference between mean IQ of PhD graduates and high school students. The use of effect-sizes highlights the importance of the magnitude of differences, which is contrary to the usual emphasis on statistical significance. Many textbooks detail how effect-sizes can be calculated from various summary statistics (e.g.,Glass, 1977; Glass, McGaw & Smith, 1981).

For example, it was possible to locate 557 studies that investigated the effects of introducing computers on students' achievement (see Hattie, 1986). Using meta-analysis, these effects can be statistically synthesized to ascertain an overall effect . The average effect-size across these 557 studies was .31.Thus, compared to classes without computers, the use of computers was associated with advancing children's achievement by approximately three months, improving the rate of learning by 15%, about 65% of the effects were positive (that is, improved achievement), thus 35% of the effects were zero or negative, and the average student achievement level after using computers exceeded 62% of the achievement levels of

the students not using computers. An effect-size of .31 would not, according to Cohen (1977), be perceptible to the naked observational eye, and would be approximately equivalent to the difference between the height of a 5'11" and a 6'0" person.

Of course, this is only an overall effect-size from introducing computers. There are many important moderators. For example, the effects decrease with age: primary students gain most (effect-size = .48), secondary students have medium gains (effect-size = .32), and college and university students gain least (effect-size = .25); there are differences in effect-sizes on achievement between males and females in secondary but not elementary classes (see Fitzgerald, Hattie, & Hughes, 1985; Hattie & Fitzgerald, 1987).

A unidimensional continuum thus has been devised that can serve to place the various effects of schooling. The scale is expressed in standard deviation units such that 1.0 is unlikely although a very obvious change in achievement, and .31 is typical after introducing computers. This continuum provides the measurement basis to address the question of the effects of schooling.

It is possible to statistically synthesize the results of a large number of studies, and thus ascertain the typical effects of schooling, and identify the innovations or changes that improve achievement in a systematically positive manner. A computer search of Psychological Abstracts, Dissertation Abstracts and ERIC identified 134 meta-analyses that related some facet of school learning to achievement outcomes - excluding studies using special class groups (full details of the studies are available in Hattie, 1987).

Altogether, 22,155 effect-sizes from 7,827 studies were computed, representing approximately 5-15 million students, and covering almost all methods of innovation. A summary of these meta-analyses is presented in Table 1.

Table 1 about here

The key question is "What is the typical effect of schooling"? The answer is derived from averaging the effects across the 134 meta-analyses and is .40.

Most innovations that are introduced in schools improved achievement by about .4 of a standard deviation. This is the benchmark figure and provides a "standard" from which to judge effects. A comparison based on typical, real-world effects rather than based on the strongest cause possible, or with the weakest cause imaginable. At minimum, this continuum provides a method for measuring the effects

of schooling.

The typical effect does not mean that merely placing a teacher in front of a class would lead to an improvement of .4 standard deviations. Some deliberate attempt to change, improve, plan, modify, and innovate is involved. As noted above the typical effect-size of .40 may not be uniform across all students as there may be many moderators.

Four overall findings

First, innovation is the theme underlying most of these effects. That is, a constant and deliberate attempt to improve the quality of learning on behalf of the system, principal and teacher typically relates to improved achievement. The implementation of innovations probably captures the enthusiasm of the teacher implementing the innovation and the excitement of the students attempting something innovative.

Second, the most powerful single moderator that enhances achievement is feedback. The most simple prescription for improving education must be "dollops of feedback". The effect-sizes for reinforcement is 1.13, remediation and feedback .65, mastery learning (which is based on feedback) .50, more specifically, homework with feedback is much more effective than homework without feedback, and recent reviews point to the power of feedback as a discriminator between more and less effective uses of computers in classrooms. This does not mean using many tests and providing over-prescriptive directions, it means providing information how and why the child understands and misunderstands, and what directions the student must take to improve.

Third, the continuum can be used to address the question as to the effects of schooling and assess the effects of teaching competencies. Over the past thirty years various researchers have provided evidence that schools and teachers are not effective in enhancing achievement (e.g., Coleman, Campbell, Hobson, McPartland, Mood, Weinfield & York,1966; Jencks, Smith, Ackland, Bane, Cohen, Gintis, Heynes & Michelson, 1972), although many have provided contrary evidence to these claims. Using the continuum of effects introduced above, the overall effect-size related to home and social influences is .38, whereas learning processes or presentation (usually controlled by the teacher) is .62; the background and style of teachers is .50; curricula differences is .48, methods of instruction is .36; and school environment (e.g., class size, environment) is .36. The student variables are .53. Thus, schooling does make a difference compared to home effects. Of more interest is the clear implication that teaching competencies make a difference. The major differences relate to certain instructional changes that teachers control and the influence that teachers have over the manner they interact with students. It is clear that within-class effects are the most dominant in enhancing achievement.

Fourth, most innovations that attempt to individualize instruction are not noted by success. Individualization programs produce an average effect-size of .14, and programmed instruction yields .18. The negligible effects of individualization are particularly important when it is recognized that students spend about 66 percent of their time working alone (Rosenshine, 1979). Too often, individualization means placing the child alone to work on a particular task, usually relating to his or her particular needs, progress, pace, and behavior. These attempts usually have little feedback, little attention by a busy teacher catering to the other 30 or so students, and the student typically has little knowledge of success or failure at the specifics of the task. The effects dramatically increase when individualization is coupled or complemented with feedback: peer tutoring and mastery learning packages are both .50. The key for the power of these individualized procedures, once again, is feedback.

Implications of the continuum

The scale can allow for more meaningful decisions as to what can be implemented to improve learning. The aim for teachers and policy-makers should not be to implement everything that enhances learning. The costs in time, energy and money would be too great. Rather, it is more desirable to compare the various costs of introducing various innovations relative to their benefits on achievement and relative to other effect-sizes along the continuum (see Levin, Glass & Meister, 1987). This is obviously true in special education classes.

It must also be emphasized that the studies reviewed have been predominantly related to achievement should not mean that schooling should predominantly relate to achievement. It is possible to construct a similar continuum related to affective outcomes. Not surprisingly, the number of studies is far less. The overall effect-size, after synthesizing 2,269 relationships between various programs and affective outcomes, was .22. The low and negative relationships included desegregation (.01), ability grouping (.16), computer-assisted instruction (.05), open versus traditional classes (.12), programmed instruction (-.14), and class size (-.09). Medium effects included remediation (.19) and co-operative learning environment (.21). Large effects on affective outcomes included peer tutoring (.34) and mastery learning (.69).

Special education

Now, the question today is how this model relates to special education.

It may surprise you but meta-analysis is alive and well in the area of special education. I was able to locate over 80 meta-analyses that specifically referred to special education.

Let me commence by providing a flavour of some of them, before I code them into my model.

Effects of early intervention

Let me use a study by Karl White to illustrate. He was interested in the efficacy of early intervention available to handicapped, disadvantaged and at-risk children. He noted that early intervention research is too often used as a political weapon or as an act of faith by practitioners. He cited the comment by one such advocate: "I use research like a drunk uses a lamp pole. I use it for support, not illumination". Of the over 2500 studies, White investigated 2266 effect-sizes (from 326 different studies) based on various research designs. The overall effect-size was .52. This represents a gain of for 8 IQ points, an improvement from the 30th to 50th percentile for motor functioning , a gain of approximately 10 months of reading for reading achievement at the second grade. Those, he claimed, were "substantial effects that are of obvious clinical importance" (p. 407).

Note the following graph that indicates the effect-sizes as a function of time at which the outcome is measured.

Graph (p. 407).

And graph for handicapped children. The immediate effect is .4, but note the small sample size. Is this a reflection of how much you care about the quality of your programs? We can surely, only conclude that those who claim that early intervention for handicapped children results in long-term impact are arguing in the absence of data (p. 408). We are coming in on a wing and a prayer - Not good enough.

As I have argued earlier, the most critical and useful benefit of meta-analysis relates to the identification of mediating variables. For example, the effects of interventions in which parents were NOT used at all or only used to a minor degree was .42 (n = 684), and the average wherein parents were used as a major or only intervener was .41 (n = 200). Similar results were obtained when data were examined regarding whether the program was centre-based, home-based, or home- and centre-based combined; whether parents or parents and children were the target of the intervention, and the degree to which the intervention program intended to involve parents (see also Tingey, 1986).

A frequently assertion is that the earlier the intervention the more effective. Note graph

(p. 411).

These data do not support the commonly held position that "earlier is better".

So what does White conclude. He commences by noting the paucity of available empirical data and concluded that "there simply is not enough information to be confident about the long-term impact of early intervention with handicapped children. There is stronger support for the immediate positive effects of intervention with disadvantaged children - which is consistent with my claim about the effects of short-term interventions (see also Fuchs & Fuchs, 1986; Tingey, 1986). Unlike White's drunk, maybe we are staying near the lamp post looking for answers, when the real answers are to be discovered far in the distance.

More on early intervention

It is necessary to note that White's meta-analysis should not be considered the end of the story. Meta-analysis does not solve the problem, it merely summarizes the earlier research, and meta-analyses differ in quality and generalizability. In White's case, when only the good quality studies were used the effects were slightly, though negligibly, larger. One of the features that attracts me to meta-analysis is that it has been used to analyze similar topics and the results, generally, are very similar. This implies a certain robustness, and this is surely confidence inspiring.

Arnold, Myette and Casto (1986) evaluated 30 studies on the same topic and reported a larger effect-size of .59, which is most similar to White's average of .52. Arnold et al. noted that significant neurological involvement appeared to inhibit the overall effects of language intervention, but other subject characteristics (such as severity of handicap, chronological age) appeared to bear little overall relation to study outcomes. Scruggs et al. (1988) performed a meta-analysis on single-subject studies of early intervention for pre-school handicapped students. There were no effects of age, sex, who presented the intervention (teachers, parents, researchers, and non-handicapped peers), but there were differences for place of instruction (schools outperforming homes), and intervention (direct instruction, time delay, reinforcement). The authors concluded that intervention effectiveness will be more closely related to environmental contingencies than to specific subject characteristics.

Other examples

To provide other illustrations, Ault, Wolery, Doyle & Gast (1989) performed a meta-analysis of comparative studies in the instruction of students with moderate and severe handicaps. Although all strategies were effective in teaching students new behaviours in at least some of the studies, the stimulus modification procedures (stimulus shaping, stimulus fading) appear to be more efficient in terms of errors and direct instruction time than trial-and-error, error correction, antecedent prompt and test, most-to-least prompting, and progressive time delay.

Casto & Mastropieri (1986) reported that the typical effect-size in intervention program that focused on parents and students was .74, and where parents of children were the focus of the intervention separately the effect-size was .44.

In studies using direct instruction, 53% favoured direct instruction, and there were no differences relating to the effectiveness relating to particular handicapping conditions, age or skill areas (White, 1988).

Fox (1989) reported that the successful generalization of skills by persons with profound mental handicaps were characterized by several trained exemplars. Trained behaviours were likely to be reinforced in natural settings by natural consequences, use of training stimuli common to the generalization setting, and training in common environments.

Schmidt et al. (1987) found that the effect-size from using with CAI was .53, and the greatest achievers were the language disordered and mentally retarded children. This is appreciably above the effect found in typical classroom (.32), but not too dissimilar from the effects from using CAI with very young students (.35). Once again, the effects are similar across special education and non-special education students. As Schmidt et al. reported, the effects of CAI were more related to moderator effects such as the amount of feedback within the CAI packages.

Tingey (1986) found that half-day programs produced as many gains as full-day programs, and programs which expected parent involvement were not more effective than those which did not. This is consistent with my earlier claim that shorter programs, which capture teachers and students enthusiasm, and thereby increasing the amount and quality of feedback is critical. Another explanation is the "little fish in big ponds" analogy. In half-day programs, special education students can be big fish in little ponds for half the day, and little fish in big ponds for the other half. In the context, I note the research that indicates that half rather than full-day programs also markedly increase student's self-concept (Strang et al, 197x).

Mastropieri & Scruggs (1986) summarized 18 studies assessing the effects of early intervention for socially withdrawn children. The overall effect-size was (r=.70). This meta-analysis highlighted the importance of direct reinforcement of the target behaviour, peer confederates were more likely to interact with target subjects when they were reinforced, and generalization was effective only when assessed immediately after the intervention in the same setting,

Mastropieri, Scruggs & Casto (1985) analyzed group studies which examined the effect of modelling on social behaviours of withdrawn pre-schoolers. The overall effect-size was .84.

Kavale & Nye (1986) used 1077 studies to assess the effects of learning disabilities in achievement, linguistic, neuropsychological and social/behavioural domains. Overall there were 6085 effect-sizes, based on 106,000 students, leading to a mean effect of .66. This indicates that approximately 75% of the LD population differs from the normal group across measures of achievement, linguistic,

neuropsychological and social/behavioural domains. The greatest effects related to the linguistic domain (.88), then achievement (.68), neuropsychological (.64), and finally social/behaviour (.58). These differences are not so large, however, such that we must conclude that language parameters are an important parameter of learning difficulty, but probably no more important than the other domains in providing a complete description of learning disabilities. There "is more than one form of LD and no uniform pattern of deficits. Consequently, unitary conceptions of LD that stress deficits in a single domain are inadequate for describing LD in general and can explain only a limited number of LD cases" (p. 457).

Skiba, Casey & Center (1986) used 315 effect-sizes (35 studies) to investigate the effects of some form of reinforcement or feedback and found an effect-size of 1.88. Unheard of improvements. The most powerful type of reinforcement was preferred activity, token, feedback and then social, but the differences were small, and all effects were substantial. Most interesting were the interactions. Social, token and activity reinforcements were clearly superior to simple feedback in group situations; but less effective than feedback when given in individual contingency arrangements. This indicates that the information function of feedback for an individual may be too diluted in a group situation to afford much control over behaviour. Reinforcement was more effective for off-task behaviour, social interactions, regular class settings; whereas feedback was more effective in the management of disruptive behaviour, and in self-contained or resource rooms.

Cook, Scruggs, Mastropieri & Casto (1986) investigated the use of handicapped students as instructional agents, or tutors. Across 74 effect-sizes (studies = 54), the overall effect was .58 for the person being tutored, and .53 for the tutor. It was more effective as a supplement rather than a substitute to teaching.

Among the more interesting meta-analyses has been Wang & Baker (1986) who investigated the effects of mainstreaming. From the few studies they located (11 studies, 115 effects) the average effect was .33. Performance effects were .44, attitudinal effects was .11, and process effects was .55. They point to the importance of the determination of the initial capabilities of individual learners, continuous assessment of learner capabilities, periodic evaluation of learn programs, student self-management, alternative routes and a variety of materials, some student choice of goals, and peer assistance and group projects. The decision to go into mainstreams thus needs to bas based on individual consultation and not on some blanket "Social Justice" edict.

Returning to the claims

The claim was that: The effects of schooling that are achieved by special education students matches that of "normal" children both in

a. the degree of effects you attain in achievement

b. the types of procedures which are effective.

Table 2 presents the overall findings from these meta-analyses relating to special education. As you can see, the effects are typically substantially greater than those relating to "normal" students. The overall effect is .60. This is greater than for "normal" students. In the language of meta-anlaysis, this difference of .2 can be regarded as small but meaningful.

The first reaction must be to reiterate an earlier statement. The effects have been achieved in special education classes by special education teachers - and NOT in normal classroom by non-specialists. Moreover, virtually all the above meta-analyses have summarized USA data and as Peter Cole has indicated there are about 6% of students in special education classes - with mainstreaming in place; and in WA there will be no more than 2% in special education classes - and we have yet to have mainstreaming. The Ministry calls this social justice whereas it is surely social stupidity.

An obvious differences is that the effect-size deliberately partials out the mean effects. Thus, I am not claiming that special education teachers achieve effects to the same degree as other teachers; only that they can improve the achievement of students by a similar amount.

Another differences may be that in special education you are aiming for more narrow outcomes based on narrow constructs. For "normal" students the comparison effects categorized as narrow constructs and narrow outcomes was .37 (sd = .11; broad constructs and narrow outcomes (M = .23, sd = .12); broad constructs and broad outcomes (M = .43, sd = .16); narrow constructs and broad outcomes (M = .35, sd = .10)).

Regarding my second claim, although there are fewer meta-analyses it appears that there are similar conclusions to the earlier study. Similar processes relate to effectiveness. Innovation and experimentation, shorter intensive programs, feedback, and the criticalness of the teacher are highlighted. As I intend to add more studies to this preliminary analyses I am sure that there will be reasonably close mapping.

I must note that so far I have been very data driven. Helen Canute and I are presently embarking on a project that assesses a model of teaching and we intend to use analyses similar to what I have outlined today to assess this model. Let me comment before I close as I do not want to leave the impression that all these data are the ultimate. Our model is based on the following:

a. That achievement is enhanced to the degree that students and teachers set challenging rather than "do your best" goals relative to the students' present competencies. Consequently,

i. there is more control over attaining challenging goals in shorter rather than longer programs

b. That achievement is enhanced to the degree that students

i. are trained to receive feedback to verify rather than enhance their sense of efficacy of achievement

ii. teachers use reinforcement to help verify rather than enhance students sense of efficacy

c. That achievement is enhanced to the extent that teachers become more automatic in many of the key teaching competencies such that they can then spend more time providing feedback to individual students.

i. it is teacher competencies rather than structural changes (e.g., mainstreaming, Ministry policy, etc) that most relates to enhanced achievement

ii. there is more teacher satisfaction in assessing whether they have attained short-term rather than long-term goals, and thus more feedback as to the efficacy of their programs.

Ask me back next year if you want more answers. For now I have one final contention.

If my major contentions today are convincing in light of the evidence I have provided then we need more attention to What Works. The biggest selling book in the history of education is "What works". Outline. Given a) that I have argued that it is what teachers do (not curriculum, programs, Ministry policies, principal behaviour) that makes the difference then you could do well to highlight "What works"; b) that few of the seemingly obvious factors that "work" are present in most classrooms then you could do well to more aware of the existence of the key factors in your classroom.

So,

we know the basis of What Works in special education;

we know that special education students can achieve as great effects as a consequence of your teaching as do "normal" students;

we know that the factors which work with "normal" students also work with special education students; and

we know that it is you as teachers that make the difference.

Descriptions of study	Authors	No. of Studies	No. of Effects	Effect -size
Efficacy of early intervention with handicapped, disadvantaged & at-risk	White	2266	326	.52
Parents not used			684	.42
Parents used			200	.41
Early intervention programs - handicapped children	Casto & Mastropieri	74		.30
Early intervention program - language	Arnold, Myette & Casto (1986)	30		.59
Early intervention	Casto & White (1984)			.50
Early intervention for socially withdrawn children	Mastropieri & Scruggs	18		.70
Early intervention of learning	Horn & Packard	58		
Language			51	1.21
Sensory			39	.93

Behavioral-emotional			17	1.09
Soft neurological			10	.92
IQ			12	1.23
Early intervention ages 2-4	Guralnick			.60
Early intervention	Rauh et al.			
low birth weight				.70
Infant Health program				.83
Intervention program - with parents	Casto & Mastropieri			.70
with parent negligible				.30
parents and students				.74
parents of children were focus				.44
Language intervention	Nye et al.	43		.35
Direct Instruction	White	25		.25
Modality learning	Kavale & Forness	39	318	.28
Auditory			80	.18
Visual			81	.09
Kinesthetic			44	.18
CAI with disabled students	Schmidt, Weinstein, Niemic & Walberg	22	48	.53
Effects of learning disabilities	Kavale & Nye	1077	6085	.66
linguistic domain				.88

achievement				.68
neuropsychological				.64
social/behaviour				.58
Reinforcement or feedback	Skiba, Casey & Center	35	315	1.88
handicapped students as tutors	Cook, Scruggs, Mastropieri & Casto	54	74	.58
those tutoring				.53
Modelling on social behaviours of	Mastropieri, Scruggs & Casto	15	44	.84
withdrawn pre-schoolers				
Mainstreaming	Wang & Baker	11	115	.33
Performance				.44
attitudinal				.11
process				.55
Overall				.60

References

Barr, R. & Dreeben, R. (1983). How Schools work. Chicago, Il: University of Chicago.

Bennett, N. (1976). Teaching styles and pupil progress. London: Open Books.

Bloom, B.S. (1976). Human characteristics and school learning. New York: McGraw-Hill.

Broudy, H.S. (1988). The uses of schooling. New York, NY: Routledge.

Cahen, S., & Davis, D. (1987). A between-grade-levels approach to the investigation of

the absolute effects of schooling on achievement. American Educational Research Journal, 24, 1-12.

Carroll, J.B. (1963). A model of school learning. Teachers College Record. 64, 722-733.

Cohen, J. (1977). Statistical power analysis for the social sciences. New York: Academic Press.

Cohen, J. (1990). Things I have learned (so far). American Psychologist, 45, 1304-1312.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F. D., & York, R.L. (1966). Equality of educational opportunity. Washington, D.C.: U.S. Office of Education.

Creemers, B.P.M., & Scheerens, J. (1989). (Eds.). Developments in School effectiveness research. International Journal of Educational Research, 13, 685-825.

Fitzgerald, D., Hattie, J.A., & Hughes, P. (1985). Computer applications in Australian classrooms. Ministry of Education, Canberra.

Glass, G.V. (1977). Integrating findings: The meta-analysis of research. In L. Shulman (Ed.), Review of research in education (Vol. 5, pp. 351-379). Itasca, IL: Peacock.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, California: Sage.

Glass, G.V. & Smith, M.L. (1979). Meta-analysis of research on class size and achievement. Educational Evaluation and Policy Analysis, 1, 2-16.

Graue, M.E., Weinstein, T., & Walberg, H.J. (1983). School-based home instruction and learning: A quantitative synthesis. Journal of Educational Research, 76, 351-360.

Hattie, J. A. (1987). Identifying the salient facets of a model of student learning: A synthesis of meta-analyses. International Journal of Educational Research, 11, 187-212.

Hattie, J.A. & Fitzgerald, D. (1987). Computers and sexism. Australian Journal of Education, 31, 3-26.

Hedges, L.V., & Olkin, I. (1985). Statistical methods for meta-analysis. New York: Academic Press.

Heyneman, S.P., & Loxley, W.A. (1983). The effects of primary-school quality on academic achievement across twenty-nine high- and low-income countries. American Journal of Sociology, 88, 1162-1194.

Jencks, C., Smith, M., Ackland, H., Bane, M.J., Cohen, D., Gintis, H., Heynes, B., & Michelson, S. (1972). Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books.

Johnson, E.G., & Zwick, R. (1990). Focusing the new design: The NAEP 1988 Technical report. The nation's report card. Princeton, NJ: ETS.

Lee, V.E. & Bryk, A.S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62, 172-192.

Levin, H.M., Glass, G.V. & Meister, G.R. (1987). A cost-effectiveness analysis of computer-assisted instruction. Evaluation Review, 11, 50-72.

Novick, M. R. & Jackson, P. H. (1974). Statistical methods for educational and psychological research. New York, NY: McGraw.

Paschal, R.A., Weinstein, T., & Walberg, H.J. (1984). The effects of homework on learning: A quantitative synthesis. Journal of Educational Research, 78, 97-104.

Raudenbush, S.W. (1988). Applications of hierarchical linear models in educational research: A review. Journal of Educational Statistics, 13, 85-116.

Raudenbush, S.W. (1989). The analysis of longitudinal, multilevel data. International Journal of Educational Research, 13, 721-740.

Rosenshine, B.V. (1979). Content, time, and direct instruction. In P. L. Peterson & H. J. Walberg (Eds.), Research on Teaching (pp. 28-56). Berkeley, California: McCutchan.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J. & Smith, A. (1979). Fifteen thousand hours. Cambridge: Harvard.

Walberg, H.J. (1981). A psychological theory of educational productivity. In F. H. Farley and N. Gordon (Eds.), Psychology and education (pp. 81-108). Berkeley, California: McCutchan.