Effectiveness of Interventions Influencing Academic Behaviors: A Quantitative Synthesis of Single-Subject Researches using the PEM Approach

Yu-Jing Gao and Hsen-Hsing Ma

The purpose of the present study is twofold: (a) to test the repeatability of the superiority of the percentage of data points exceeding the median of the baseline phase (PEM) approach over the percentage of nonoverlapping data (PND) approach for the synthetic analysis of single-subject researches, and (b) to demonstrate the application of the PEM approach in conducting a quantitative synthesis of single-subject researches, which investigated the effectiveness of interventions on academic behaviors. The analyzed studies were obtained through a computer-assisted search of the relevant databases and a hand search of the relevant behavior analysis journals. The major finding demonstrates the repeatability of the superiority of the PEM approach over that of the PND.
Key words: PEM approach (the percentage of data points exceeding the median of the baseline phase); PND approach (the percentage of nonoverlapping data); Synthetic analysis of single-subject researches.

The purpose of the present study is twofold: (a) to test the repeatability of the superiority of the percentage of data points exceeding the median of baseline phase (PEM) approach (Ma, 2006) over the percentage of nonoverlapping data (PND) approach (Mastropieri & Scruggs, 1985-86) for the synthetic analysis of single-subject researches, and (b) to demonstrate the application of the PEM approach in conducting a quantitative synthesis of single-subject researches, which investigated the effectiveness of interventions on academic behaviors. Ma's (2006) study showed that PEM approach had a higher validity than the PND approach with respect to the intercorrelations of their effect size scores with original authors' judgments as well as the closeness of means of effect size in each of three categories of original authors' judgment (highly moderately and not effective) to the criteria set by Scruggs, Mastropieri, Cook, & Escobar (1986), i.e., PEM approach had higher rank correlations with original authors' judgments than PND approach did, and Mean effect size of the PEM approach in each of the categories of effectiveness fell in the range set by Scruggs, et al. (1986), but that of the PND approach fell outside of the range. Ma (2006) described and discussed the methodology of PEM approach in detail. Whether his findings are repeatable has to be tested.

The second purpose of the present study therefore is to use the PEM approach to conduct a systhetic analysis of the effectiveness of interventions on academic behaviors. Numerous educational and psychological researchers have used single-subject experimental designs to evaluate the effectiveness of interventions intended to promote academic behaviors. It is therefore meaningful to conduct a synthetic analysis to determine whether these intervention strategies are effective and which one shows a larger effectiveness. Through a review of the literature, it is possible to identify several intervention strategies which have been adopted by behavior analysts to enhance academic behaviors, such as computer-assisted instruction (Higgins and Boone, 1990; Howell, Sidorenko, & Jurica, 1987); cooperative learning (Pigott, Fantuzzo, & Clement, 1986; Olympia, Sheridan, Jenson, & Andrews, 1994; Cushing & Kennedy, 1997); mastery learning (McDowell & Keenan, 2001); reinforcement (Noell, et al. 1998; Lloyd, Eberhardt, & Drake, 1996; Gillat & Sulzer & Azaroff, 1994); self-control training (Stevenson & Fantuzzo, 1986; Harris, 1986; Robert, Nelson, & Olson, 1987; Dunlap & Dunlap, 1989; Lloyd, Bateman, Landrum, & Hallahan, 1989); social learning (Stevenson & Fantuzzo; 1984); instruction in learning strategies including instruction in cognitive and metacognative strategies (Montague, 1992), training in phonological awareness training and in word analogies (O'Shaughnessy & Swanson, 2000), story mapping technique (Babyak, Koorland & Mathes, 2000; Gardill & Jitendra, 1999; Gurney, Gersten, Dimino, & Carnine 1990; Idol, 1987; Newby, Caldwell, & Recht, 1989), and training in the question-asking (Knapczyk, 1989). These intervention strategies were found to be positively effective by the respective authors. Which ones are more effective? The PEM approach will be used to address this question.

Method

Procedures for locating studies

The single-subject researches on academic behaviors used in this synthesis were obtained through a computer-assisted search of the relevant databases, including ERIC, EBSCOHost, ProQuest, and PSYINFO. Descriptors included academic behavior, single-subject, and behavioral modification. A hand search of relevant behavior analysis journals was also conducted in journals such as Behavior Modification (1984-2003), Behavior Therapy (1977-2003), Behavioral Disorders (1981-2003), Bulletin of Special Education (1985-2003, in Chinese), Bulletin of Special Education and Rehabilitation (1991-2001, in Chinese), Journal of Applied Behavior Analysis (1968-2003), Journal of Learning Disabilities (1975-2003), Journal of Special Education (1986-2002, in Chinese), Learning Disability Quarterly (1981-2003), and The Journal of Special Education (1967-2003). The lists of references in the studies found by the above-mentioned methods were traced to discover remaining usable studies not yet identified. Studies that meet the following criteria were included in this synthesis: (a) the intervention (independent variable) focused on promoting academic behaviors (dependent variable), (b) a valid and scientific single-subject research design such as reversal (withdrawal) or multiple-baseline design was employed, (c) graphic time-series data displays suitable for calculating PEM and PND were provided. Studies, which included participants with mental retardation, developmental disabilities, or physical handicap, were excluded.

Procedure for coding the Study

Study characteristics. Variables in each of the following areas were coded:

1. Authors' conclusion on the overall effectiveness of treatment (2 = was highly effective, participants improved immediately, showed substantial improvement, progressed, showed a positive shift, showed gradual increase, showed a meaningful increase, attained a high level; 1

= was partially effective, was moderately effective, slight decrease, increased but had data overlapping, increased but variably, showed a small increase; 0 = had questionable or no effect, produced little improvement, had minimal effect, virtually unchanged, the treatment had no noticeable effect).

2. Categorization of dependent variables. Dependent variables were classified into two categories: academic achievement and academic engagement. Academic achievement is related to the formally-noted ability of the participants (i.e., the percentage of accurate scores, number of problems completed correctly, and grades, etc.) and academic engagement is related to the involvement of the participants in academic tasks or school work including task completion, degree of attention, on-task behavior, the following of the teacher's directions or instruction, participation in class discussion, and asking the teacher for assistance if needed.

3. Categorization of independent variables. Independent variables were divided into eight categories: (a) computer-assisted instruction including drill and practice software and tutorial-based software, (b) co-operative learning (students in a small group performed separately the role of peer instruction, peer observation, peer evaluation or peer reinforcement), (c) mastery learning (students practiced repeatedly until they met a specified criterion), (d) reinforcement (primary and/or secondary reinforcers were contingent on the accuracy and/or amount of tasks completed), (e) self-control training (the components of self-instruction, self-monitoring, self-evaluation, and self-reinforcement were trained in a separate or combined form), (f) social learning (students learning of behavior patterns through watching the actions of others and observing the consequences), (g) learning strategy instruction including cognitive and metacognitive strategy instruction, and reading comprehension training (training in word recognition, phonological awareness, analogy, and story mapping techniques), and (h) others including student-selected seatwork assignment, extinction of teacher attention, public posting, student-operated business curriculum, and functional writing.

4. 4. Participants' age or educational level. Age or educational level were classified by five groups: preschool, elementary school, middle school, high school, and college.

5. Participant classification. Participants were classified as manifesting attention deficit hyperactivity disorder, autism, behavioral disorder, emotional disturbance, learning disabilities, multiple handicapped (participants having two or more above-mentioned diagnoses), and normal.

6. Subject matters. Subject matters in which instructions were given included educational psychology, language, mathematics, science, and social science.
7. Setting. Intervention settings were classified as home, institution, school and others.

8. Interveners. Interveners who carried out the treatment were categorized into parent, peer, principal, researcher (experimenter), and teacher.

9. Order of pairs of baseline-treatment phase of a reversal design: 1 = the first pair, 2 = the second pair. Each pair of baseline-treatment phases in a multiple baseline design was treated as the first pair except in the case that it contained a reversal design.

10. Type of experimental designs: 1 = reversal design, 2 = multiple design.

Computation of effect sizes

Calculation of the PEM and PND scores. To compute the PEM scores, the first step is to draw a horizontal median line in the baseline phase. This line will hit the median when the number of data points in the baseline phase is odd and go between the two middle points if the number of data points is even. The median line will stretch out horizontally to the treatment phase. The second step is to calculate the percentage of the data points of the treatment phase above the median line. If the treatment is to decrease an undesirable behavior, then the PEM score will be the percentage of data points below the median line in the treatment phase. A PND score is computed by calculating the proportion of data points in a treatment phase above the highest data point of the immediately preceding baseline phase (or under the lowest data point based on the hypothetically expected direction).

Reliability. The first author of the present study chiefly carried out the coding work. Another doctoral student was asked to calculate independently the scores of the PND scores and the PEM. The percentage of agreement between two raters was calculated by the formula: the number of agreements divided by the number of agreements and disagreements. Disagreements were resolved by discussion and re-calculation.

Results

A total of 98 studies that investigated the effectiveness of interventions on academic behaviors met the coding criteria and were included in this synthetic analysis. The mean of the PEM scores of 952 effect sizes from the 98 studies was .87 with a standard deviation of .25, which demonstrated a moderate effectiveness of the treatments according to the criteria set by Scruggs, Mastropieri, Cook, & Escobar (1986). They suggested that scores of .9 and higher represent highly effective outcomes, scores of .7 to .9 represent fair outcomes, scores of .5 to .7 represent questionable effect and scores below .5 represent basically unreliable treatments. As it is hard to distinguish questionable and unreliable effects by a visual judgment of the graphic display, such effects were pooled together and classified as having no effect on the reporting of the results in the present study.

The question then arises as to whether or not a mean effect size of .87 is significantly different from .5, the null hypothesis of PEM scores. In order to be tested with parametric statistics, such as a t-test, the data must meet the assumptions of such statistical tests with regardd to the normality, homogeneity, and independence of the distribution of residuals (Myers, 1972, p.61). In order to test whether the residuals were independently distributed, the residuals of 952 effect sizes were created by using " center" in the ARIMA (autoregressive integrated moving average) procedure to subtract each effect size from the mean effect size (SAS Institute Inc. 1984, p.131). Lag 1 of autocorrelation function was found to be .44 with a standard error of .03, $p < .001$. This result indicated that the assumption of the independent distribution of the residuals was violated and that parametric statistics were not for the testing of whether or not the mean effect size of .87 is significantly different from .5. However, when the effect sizes of a study were averaged and treated as a unit of analysis, the distribution of residuals would be independent, because each mean effect size came from different article. Using the ARIMA procedure mentioned above to produce the residuals of the 98 averaged effect sizes and then Lag 1 of the autocorrelation function of the 98 residuals was found to be .13 with a standard error of .10, $p > .05$. This result denotes that the assumption of the independence of the residuals was met. The grand mean of the 98 independent mean effect sizes was .88 with a standard deviation of .13. The results of a one-sample t-test, $t(97) = 28.79$, $p < .001$ showed that the intervention strategies had a significant influence on academic behaviors. In order to make comparisons with the scores produced by the PND approach, the PND scores, as well as the coding number of the judgments of the original authors' are displayed side by side with the PEM scores throughout the present study.

Reliability

To calculate the reliability of the coding procedure, 25 studies were selected as a random sample from the 98 studies by the use of a table of random numbers. The reliability between the raters for all of the coded variables was 94.45%. In addition, the reliability of the coding of the PEM scores, the PND scores, and the ratings of judgments of the original authors were 96.17%, 97.13%, and 89.95% respectively.

Validity

The Spearman rank order correlation coefficient was used as the index of validity because the rating from the judgments of the original authors is coded with an ordinal scale.

TABLE 1
Intercorrelation Between the Judgments of Original Authors', the PND scores, and the PEM scores

	Judgments of the original authors	PND scores	PEM scores
Judgments of the original authors	_	.47*** (N = 937)	.59*** (N = 937)
PND scores	.44*** (N = 98)	_	.63*** (N = 952)
PEM scores	.61*** (N = 98)	.73*** (N = 98)	_

Note. The correlation coefficients between the PEM and PND scores are Pearson r because both the PEM and PND scores are on an interval scale while other correlation coefficients are Spearman correlation coefficient because the ratings of the original authors' judgments are coded with an ordinal scale.
*** $p < .001$

As shown in Table 1, the values under the diagonal represent the intercorrelations among the PEM scores, PND scores, and ratings of the judgments of the original authors with a single article as the unit of calculation, whereas the values above the diagonal represent their intercorrelations with a single pair of baseline-treatment as the unit of calculation. The results show that both the PEM and PND scores correlate with the ratings of judgments of the original authors significantly and that the correlation coefficient between the PEM scores and the conclusions reached by the original authors was higher than that between the PND scores and the original authors' judgments.

The Influence of Orthogonal Slope Changes and Outliers

Only four studies out of 154 ABAB designs were found to display orthogonal slope changes in the second pair of baseline-treatment phases (Hasazi & Hasazi, 1972; Howell, Sidorenko, & Jurica, 1987; Lahey, McNees, & Brown, 1973; and Olympia, Sheridan, Jenson, & Andrews, 1994). This finding indicates that the amount of orthogonal slope changes that appeared in the second pair of baseline-treatment phases was not large enough to underestimate the effect size in any threatening way.

Sixty-seven of 952 effect sizes were underestimated by the PND approach owing to the influence of the outliers that appeared in the baseline. After excluding these 67 effect sizes, the Spearman rank correlation coefficient between the PND scores and original authors' judgments was .67, as shown in Table 2. It was higher than the previous result (r= .47) as shown in Table 1, but still lower than the correlation coefficient between PEM scores and the judgments of the original authors.

TABLE 2
Intercorrelation Between the Judgments of Original Authors', the PND Scores, and the PEM Scores with 67 Effect Sizes Containing Outliers Excluded

	Judgments of the original authors	PND Scores	PEM Scores
Judgments of the original authors	_	.67*** (N=870)	.78*** (N=870)
PND Scores	_	_	.72*** (N=885)

Note. The correlation coefficients between the PEM and PND scores are Pearson r because both the PEM and PND scores are on an interval scale while other correlation coefficients are Spearman correlation coefficient because the ratings of the original authors' judgments are coded with an ordinal scale.
*** $p < .001$

Judgment Criteria of Effectiveness

The judgments of the original authors in regard to the effectiveness of treatments were classified into three categories: 'highly effective', 'moderately effective', and 'not effective' outcomes. Then the PEM and PND scores within each category were averaged. The results are exhibited in Table 3.

Table 3
Comparisons of Means of PEM and PND Scores with Criteria Suggested by Scruggs, et al. (1986) at Each Level of Effectiveness Judged by Original Aauthors

Judgments of the original authors	N	PEM	PND	By criterion of Scruggs, et al. (1986)
Highly effective	810	.94	.75	? .9
Moderately effective	26	.79	.49	? .7 < .9
Questionable or not effective	101	.35	.12	<. 7

It was found that the means of the PEM scores in each category of effectiveness were located within the range set by Scruggs, et al. (1986) whereas the means of the PND scores in each category were underestimated.

Mean Effect Sizes of Dependent and Independent Variables

The mean effect sizes of each category of academic behaviors and intervention strategies are presented in Table 4.

Table 4
Mean Effect Sizes of Dependent and Independent Variables

	PEM				PND		Judgments of the original authors		
	N	Mr[a]	M	SD	M	SD	N	M	SD
Overall effect									
With article as unit	98		.88	.13	.67	.24	98	1.79	.40
With pairs of baseline-treatment phases as unit	952		.87	.25	.67	.38	937	1.76	.63
Dependent variables									
Engagement	237	517	.89	.24	.69	.38	237	1.73	.67
Achievement	715	463	.86	.25	.66	.38	700	1.76	.62
Independent variables									
1. Computer assisted instruction	11	264	.76	.20	.44	.29	11	2.00	.00
2. Co-operative learning	137	445	.84	.26	.59	.41	124	1.76	.65
3. Mastery learning	37	527	.92	.20	.86	.25	37	2.00	.00
Table 4 (continued)									
4. Reinforcement	162	515	.92	.19	.70	.36	162	1.83	.53
5. Self-control training	303	520	.90	.23	.70	.37	303	1.83	.54
6. Social learning	40	472	.85	.25	.64	.36	40	1.70	.72
7. Instruction in learning strategies	226	419	.80	.30	.64	.37	224	1.62	.76
8. Others	36	426	.83	.25	.62	.41	36	1.44	.88

Note. The numbers of PND scores are the same as the numbers of the PEM scores (N); By the multiple post hoc comparisons of different independent variables using the Mann-Whitney U test, the results showed as follows: (2, 3, 4, 5, 6) > 1; (4, 5) > (2; 7,8); 3 > 7.The numbers represent the serial number of each independent variable in the table. The numbers in the parentheses refer to the fact that the mean effect size of these variables are all significantly larger than that of the variable(s) behind the ">".
[a]Mr = mean of ranks

Throughout the following analyses, the pairs of baseline-treatment phases will be used as the unit of analysis. Because the assumption of the independence of the distribution of the residuals was violated, only a nonparametric statistical test, Kruskal-Wallis analysis of variance by ranks was used to test the significance of the difference of in the mean ranks between groups of data. Mann-Whitney U test was used to make multiple post hoc comparisons among different mean ranks.

The dependent variables were aggregated into two categories (i.e., academic engagement and academic achievement). The mean rank of effect size of engagement was higher than that of academic achievement. The Mann-Whitney U test resulted in Z (N= 952) = -3.097, $p < .002$. This finding means that it is easier to intervene to aid students in completing a task than to improve the correctness of tasks done by students.

An analysis of the PEM scores revealed that all independent variables had a positive effect on academic behaviors. The results in able 4 indicate that mastery learning, reinforcement, and self-control training were highly effective whereas the remaining interventions had only moderate effect on the improvement of academic behaviors. The Kruskal-Wallis one-way ANOVA by ranks test showed that the main effect was significant, $?^2$ (7, N = 952) = 44.24, P < .01. The results of multiple post hoc comparisons using the Mann-Whitney U test are summarized in the note for Table 4. For instance, "(2, 3, 4, 5, 6) > 1" means that the mean rank of co-operative learning (2), mastery learning (3), reinforcement (4), self-control training (5), and social learning (6) were significantly higher than that of computer assistant instruction (1), i.e., the effectiveness of the former independent variables were all larger than that of the later. Normally it is the case that the larger the mean effect size, the larger the mean rank of effect size. However, there are some inconsistencies in the rank orders of M (mean) and Mr (mean of ranks) in each subcategory of independent variables and moderators. The phenomenon is caused by the heterogeneity of the variance of the residuals and outliers of the effect sizes. It is similar to the fact that in the post hoc comparisons, the use of different parametric statistics, such as Scheffé and Duncan, would result in different conclusions of in the test of significance.

Mean Effect Sizes of Moderators

Table 5 provides information on the mean and mean ranks of the effect sizes of the moderators, i.e., the study characteristics.

Table 5

Mean Effect Sizes by Study Characteristics

	PEM				PND		Authors' judgment		
	N	Mr [a]	M	SD	M	SD	N	M	SD
Sex									
1. Female	102	271	.90	.19	.71	.34	101	1.86	.49
2. Male	439	268	.88	.22	.67	.37	436	1.79	.60
Educational Levels									
1. Preschool	9	318	.8	.21	.63	.33	9	1.56	.88
2. Elementary school	609	466	.88	.23	.7	.36	605	1.77	.61
3. Middle school	210	472	.86	.28	.65	.39	210	1.73	.68
4. High school	80	453	.89	.19	.6	.39	80	1.9	.38
5. College	16	346	.69	.38	.23	.4	5	2	0
Participants									
1. Attention deficit hyperactivity disorder	27	363	.76	.32	.63	.37	25	1.48	.77
2. Multiple handicaps	51	419	.83	.28	.55	.41	51	1.54	.83
3. Autism	16	381	.83	.26	.52	.40	16	1.88	.50
4. Behavioral disorder	145	486	.88	.25	.67	.39	145	1.74	.66
5. Emotional disturbance	15	549	.90	.28	.85	.31	15	1.87	.52
6. Learning disability	556	468	.87	.25	.69	.37	554	1.76	.63
7. Normal	120	475	.88	.24	.64	.40	109	1.85	.49
Courses									
1. Psychology	16	330	.69	.37	.23	.4	5	2	0
2. Language	383	407	.85	.26	.63	.38	379	1.72	.67
3. Math	430	452	.87	.26	.71	.37	430	1.74	.65
4. Science	12	582	1	0	.88	.25	12	2	0
5. Social Science	20	434	.88	.19	.68	.39	20	1.9	.45
Settings									
1. Home	60	600	.97	.15	.84	.25	60	1.98	.13
2. Institution	110	431	.82	.29	.66	.38	110	1.62	.77
3. School	770	475	.87	.25	.66	.38	755	1.77	.62
4. Others	12	381	.76	.32	.53	.42	12	1.33	.98
Agents									
1. Parent	23	587	.98	.07	.78	.27	23	1.96	.21
2. Peer	131	436	.84	.26	.57	.41	118	1.75	.67
3. Principal	34	576	.98	.05	.79	.33	34	2.00	.00
4. Researcher	282	452	.85	.26	.66	.37	282	1.70	.67
5. Teacher	482	490	.87	.26	.68	.38	480	1.76	.63
Designs									

1. Multiple baseline	585	466	.86	.27	.66	.37	583	1.73	.65
2. Reversal	367	491	.88	.23	.68	.38	354	1.80	.59

Note. The numbers of the PND scores are the same as the numbers of the PEM scores; By the multiple post hoc comparisons using the Mann-Whitney U test, the results showed as follows: Participants: (4, 5, 6, 7) > 1; course: 4 > 3 > (1, 2); setting; 1 > (2, 3, 4); agents: 1 > (2, 4, 5), (3, 5) > 2, 3 > (4, 5), 5 > 4;
[a]Mr = mean of ranks

There was no significant difference in the mean ranks of the effect sizes of the independent variables with regard to possible differences in the academic behaviors between female and male. The result of the Mann-Whitney U test showed that Z (N = 541) = -.09, p = .93. There was also no significant difference in the mean ranks of the effect sizes between different educational levels. The Kruskal-Wallis one-way ANOVA by ranks test resulted in $?^2$ (4, N = 924) = 8.72, p = .069. This result means the effectiveness of the independent variables on the academic behaviors will not be influenced by the sex or education levels of the participants (subjects) in an experiment.

However certain other characteristics of participants could significantly affect the effectiveness of intervention on the academic behaviors. The Kruskal-Wallis one-way ANOVA by ranks test showed that $?^2$ (6, N = 930) = 13.35, p = .04. The Mann-Whitney U tests showed that the effectiveness of intervention were with regard to those participants who were diagnosed as manifesting behavior disorder, emotional disturbance, learning disability, and as normal larger than by Attention Deficit Hyperactivity Disorder. In the present study, about 60% of participants were diagnosed as having a learning disability.

The effectiveness of interventions on academic behaviors also depended on the nature of the courses in which the interventions were carried out. The Kruskal-Wallis one-way ANOVA by ranks test demonstrated that $?^2$ (4, N = 861) = 18.68, p = .001. The Mann-Whitney test showed that the effectiveness of an intervention applied in a science course was larger than in one for mathematics, which in turn was larger than in courses of psychology and language.

The kind of agents (interveners) who implemented an intervention had an impact on the effectiveness of an intervention on academic behaviors. The Kruskal-Wallis one-way ANOVA by ranks test revealed that $?^2$ (4, N = 952) = 20.09, p < .001. It is worth noting that an intervention had the largest effect if parents carried it out. Parallel to this finding was that a home location was the most effective place to implement an intervention for the improvement of students' academic behaviors. The Kruskal-Wallis one-way ANOVA by ranks test showed that $?^2$ (3, N = 952) = 23.33, p < .001.

The impact of the kind of experimental design was not significant.
The Mann-Whitney U test resulted in Z (N = 952) = -1.50, p > .05.

Discussion

The major finding of the present study demonstrates that the use of the PEM approach is appropriate for a quantitative synthesis of single-subject researches. The PEM scores correlate more closely with the original authors' judgments in regard to the effectiveness of the various treatments than the use of the PND approach does. There are only four studies that displayed orthogonal slope changes in the second pair of baseline-treatment phases. This finding confirms that of Scruggs, et al. (1985-1986). The various kinds of interventions show moderate to highly effects on participants' academic behavior. Of all the interventions, mastery learning, reinforcement, and self-control training showed a more notable effectiveness.

The results of the present study have confirmed similar findings in the following previous studies employing between-group researches as in the studies: by (a) Guskey and Pigott (1988) as well as Kulik, et al. (1990) where it was found that mastery learning strategies had a positive effects on students' academic achievement, (b) by Lysakowski and Walberg (1981) where it was found that reinforcement had a strong effect on learning, (c) by Fletcher-Flinn and Graratl (1995) and Schmidt, et al. (1985-1986) where it was discovered that computer assisted instruction had a significant but also relatively smaller effect (mean effect size d = 0.24) on academic performance, and (d) by Fan and Chen's (2001) where it was found that parental involvement had a positive influence on students' academic achievement in a home settings and when participants' parents acted as agents to enhance academic behaviors.

A weakness of the present study is that unpublished studies were not included because of the difficulty in collecting them. Separately, although the superiority the PEM approach over the PND approach was evidenced in the present study, however, the contribution of the PND approach should not be neglected. The contribution of Mastropieri and Scruggs (1985-86) to the synthetic analysis of single-subject researches is as great as that of Glass (1976) to the meta-analysis of between-group researches. The Authors of the present study hope that a new approach to the synthetic analysis of single-subject researches will emerge in the future that can additionally address the problems of insensitivity to the magnitude, trend, and variability of data points above the median in the PEM approach as mentioned by Ma (2006).

REFERENCES

References marked with asterisk indicate studies included in the meta-analysis.

*Ayllon, T., & Roberts, M. D. (1974). Eliminating discipline problems by strengthening academic performance. *Journal of Applied Behavior Analysis, 7,* 71-76.

*Babyak, A. E., Koorland, M., & Mathes, P. G. (2000). The effect of story mapping instruction on reading comprehension of students with behavioral disorders. *Behavioral Disorders, 25(3),* 239-258.

*Billingsley, F. F. (1977). The effects of self- and externally- imposed schedules of reinforcement on oral reading performance. *Journal of Learning Disabilities, 10(9),* 20-30.

*Blankenship, C. S., & Baumgartner, M. D. (1982). Programming generalization of computational skills. *Learning Disability Quarterly, 5*, 152-162.

*Blick, D. W., & Test, D. W. (1987). Effects of self-recording on high-school students' on task behavior. *Learning Disability Quarterly, 10*, 203-213.

*Bryant, L. E., & Budd, K. S. (1982). Self-instructional training to increase independent work performance in preschoolers. *Journal of Applied Behavior Analysis, 15(2*), 259-271.

*Campbell, B. J., Brady, M. P., & Linehan, S. (1991). Effects of peer-mediated instruction on the acquisition and generalization of written capitalization skills. *Journal of Learning Disabilities, 24(1),* 6-14.

*Carr, S. C., & Punzo, R. P. (1993). The effects of self-monitoring of academic accuracy and productivity on the performance of students with behavioral disorders. *Behavioral Disorders, 18(4),* 241-250.

*Case, L. P., Harris, K. R., & Graham, S. (1992). Improving the mathematical problem-solving skills of students with learning disabilities: Self-regulated strategy development. *The Journal of Special Education, 26(1)*, 1-19.

*Chang, Y. P. (1993). The effect of reinforcement strategy in a computer assisted instruction program in the learning of addition for learning disabled children. *Bulletin of Special Education and Rehabilitation, 3,* 39-68. (National Tainan Teachers College, Tainan, Taiwan, R.O.C.) (In Chinese)

*Chou, T. J., & Jan, W. H. (1995). The effect of the metacognitive reading strategy instruction on the reading comprehension ability of elementary school children with reading disabilities. *Bulletin of Special Education and Rehabilitation, 4*, 109-152. . (In Chinese)

*Chou, T. J., & Lin, Y. H. (1996). The effects of self-instructional training on ADHD children. *Journal of Special Education, 11*, 239-284. (In Chinese)

*Christie, D. J., Hiss, M., & Lozanoff, B. (1984). Modification of inattentive classroom behavior. *Behavior Modification, 8(3),* 391-406.

*Clark, F. L., Deshler, D. D., Schumaker, J. B., Alley, G. R., & Warner, M. M. (1984). Visual imagery and self-questioning: Strategies to improve comprehension of written material. *Journal of Learning Disabilities, 17(3),* 145-182.

*Copeland, R. E., Brown, R. E., & Hall, R. V. (1974). The effect of principal-implemented techniques on the behavior of pupils. *Journal of Applied Behavior Analysis, 7,* 77-86.

*Coyne, P. D. (1978). The effects of peer tutoring with group contingencies on the academic performance of college students. *Journal of Applied Behavior Analysis, 11(2*), 305-307.

*Cushing, L. S., & Kennedy, C. H. (1997). Academic effects of providing peer support in general education classrooms on students without disabilities. *Journal of Applied Behavior Analysis, 30*, 139-151.

*Davis, R. W., & Hajicek, J. O. (1985). Effects of self-instructional training and strategy training on a mathematics task with severely behaviorally disordered students. *Behavioral Disorders, 10(4),* 275-282.

*De La Paz, S. (1997). Strategy instruction in planning: Teaching students with learning and writing disabilities to compose persuasive and expository essays. *Learning Disability Quarterly, 20*, 227-248.

*Dugan, E., Kamps, D., Leonard, B., Watkins, N., Rheinberger, A., & Stackhaus, J. (1995). Effects of cooperative learning groups during social studies with autism and fourth-grade peers. *Journal of Applied Behavior Analysis, 28,* 175-188.

Dunlap, L. K., & Dunlap, G. (1989). A self-monitoring package for teaching subtraction with regrouping to students with learning disabilities. *Journal of Applied Behavior Analysis, 22(3),* 309-314.

Fan, X., & Chen, M. (2001). Parental involvement and student's academic achievement: A meta-analysis. *Educational Psychology Review, 13(1),* 1-22.

*Fantuzzo, J. W., & Clement, P. W. (1981). Generalization of the effects of teacher- and self-administered token reinforcers to nontreated students. *Journal of Applied Behavior Analysis, 14(4),* 435-447.

Fletacher-Flinn, C. M., & Gravatt, B. (1995). The efficacy of computer assisted instruction (CAI): A meta-analysis. *Journal of Educational Computing Research, 12(3),* 219-242.

*Freeland, J. T., & Noell, G. H. (1999). Maintaining accurate math responses in elementary school students: The effects of delayed intermittent reinforcement and programming common stimuli. *Journal of Applied Behavior Analysis, 32*, 211-215.

*Freeman, T. J., & McLaughlin, T. F. (1984). Effects of a taped-words treatment procedure on learning disabled students' sight-word oral reading. *Learning Disability Quarterly, 7*, 49-54.

*Gardill, M. G., & Jitendra, A. K. (1999). Advanced story map instruction: Effects on the reading comprehension of students with learning disabilities. *The Journal of Special Education, 33(1),* 2-17.28.

*Gillat, A., & Sulzer-Azaroff, B. (1994). Promoting principals' managerial involvement in instructional improvement. *Journal of Applied Behavior Analysis, 27*, 115-129.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.

*Glomb, N., & West, R. P. (1990). Teaching behaviorally disordered adolescents to use self-management skills for improving the completeness, accuracy, and neatness of creative writing homework assignments. *Behavioral Disorders, 15(4),* 233-242.

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

*Gurney, D., Gersten, R., Dimino, J., & Carnine, D. (1990). Story grammar: Effective literature instruction for high school students with learning disabilities. *Journal of Learning Disabilities, 23(6),* 335-342.

Guskey, T. R., & Pigott, T. D. (1988). Research on group-based mastery learning programs: A meta-analysis. *Journal of Educational Research, 81(4),* 197-216.

*Hall, R. V., Lund, D., & Jackson, D. (1968). Effects of teacher attention on study behavior. *Journal of Applied Behavior Analysis, 1,* 1-12.

*Hall, R. V., Panyan, M., Rabon, D., & Broden, M. (1968). Instructing beginning teachers in reinforcement procedures, which improve classroom control. *Journal of Applied Behavior Analysis, 1,* 315-322.

*Hallahan, D. P., Marshall, K. J., & Lloyd, J. W. (1981). Self-recording during group instruction: Effects on attention to task. *Learning Disability Quarterly, 4*, 407-413.

*Harris, K. R. (1986). Self-monitoring of attentional behavior versus self-monitoring of productivity: Effects on on-task behavior and academic response rate among learning disabled children. *Journal of Applied Behavior Analysis, 19*, 417-423.

*Harris, K. R., & Graham, S. (1985). Improving learning disabled students' composition skills: Self-control strategy training. *Learning Disability Quarterly, 8*, 27-36.

*Harris, K. R., Graham, S., Reid, R., McElroy, K., & Hamby, R. S. (1994). Self-monitoring of attention versus self-monitoring of performance: Replication and cross-task comparison studies. *Learning Disability Quarterly, 17,* 121-139.

*Hasazi, J. E., & Hasazi, S. E. (1972). Effects of teacher attention on digit-reversal behavior in an elementary school child. *Journal of Applied Behavior Analysis, 5(2)*, 157-162.

*Hay, W. M., Hay, L. R., & Nelson, R. O. (1977). Direct and collateral changes in on-task and academic behavior resulting from on-task versus academic contingencies. *Behavior Therapy, 8,* 431-441.

*Higgins, K., & Boone, R. (1990). Hypertext computer study guides and the social studies achievement of students with learning disabilities, remedial students and regular education students. *Journal of Learning Disabilities, 23(9),* 529-540.

*Houten, R. V., & Houten, J. V. (1977). The performance feedback system in the special education classroom: An analysis of public posting and peer comments. *Behavior Therapy, 8,* 366-376.

*Houten, R. V., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis, 8,* 449-457.

*Howell, R., Sidorenko, E., & Jurica, J. (1987). The effects of computer use on the acquisition of multiplication facts by a student with learning disabilities. *Journal of Learning Disabilities, 20(6),* 336-340.

*Humphrey, L. L., & Karoly, P. (1978). Self-management in the classroom: Self-imposed response cost versus self-reward. *Behavior Therapy, 9,* 592-601.

*Idol, L. (1987). Group story mapping: A comprehension strategy for both skilled and unskilled readers. *Journal of Learning Disabilities, 20(4),* 196-204.

*Jolivette, K., Wehby, J. H., Canale, J., & Massey, N. G. (2001). Effects of choice-making opportunities on behavior of students with emotional and behavioral disorders. *Behavioral Disorders, 26(2),* 131-145.

*Joseph, L., & Cooper, J. O. (1991). Fourth-grade students' math performance with the criterion referenced curriculum. *Behavior Modification, 15(2*), 228-249.

*Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. (1994). Classwide peer tutoring: An integration strategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis, 27*, 49-61.

*Kamps, D. M., Leonard, B., Potucek, J., & Garrison-Harrell, L. (1995). Cooperative learning groups in reading: An integration strategy for students with autism and general classroom peers. *Behavioral Disorders, 21(1),* 89-109.

*Kelley, M. L., & Stokes, T. F. (1982). Contingency contracting with disadvantaged youths: Improving classroom performance. *Journal of Applied Behavior Analysis, 15(3),* 447-454.

*Kern, L., Childs, K. E., Dunlap, G., Clarke, S., & Falk, G. D. (1994). Using assessment-based curricular intervention to improve the classroom behavior of a student with emotional and behavioral challenges. *Journal of Applied Behavior Analysis, 27*, 7-19.

*Kirby, F. D., & Shields, F. (1972). Modification of arithmetic response rate and attending behavior in a seventh-grade student. *Journal of Applied Behavior Analysis, 5*, 79-84.

*Kirby, K. C., Holborn, S. W., & Bushby, H. T. (1981). Word game bingo: A behavioral treatment package for improving textual responding to sight words. *Journal of Applied Behavior Analysis, 14*, 317-326.

*Knapczyk, D. P. (1989). Generalization of student question asking from special class to regular class setting. *Journal of Applied Behavior Analysis, 22,* 77-83.

*Knapczyk, D. R., & Livingston, G. (1973). Self-recording and student teacher supervision: Variable within a token economy structure. *Journal of Applied Behavior Analysis, 6,* 481-486.

*Knapczyk, D .R., & Livingston, G. (1974). The effects of prompting question-asking upon on-task behavior and reading comprehension. *Journal of Applied Behavior Analysis, 7*, 115-121.

*Kraetsch, G. A. (1981). The effects of oral instructions and training on the expansion of written language. *Learning Disability Quarterly, 4,* 82-90.

Kulik, C. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational research, 60(2),* 265-299.

*Lahey, B. B., McNees, M. P., & Brown, C. C. (1973). Modification of deficits in reading for comprehension. *Journal of Applied Behavior Analysis, 6,* 475-480.

*Lane, K. L., O'Shaughnessy, T. E., Lambros, K. M., Gresham, F. M., & Beebe-Frankenberger, M. E. (2001). The efficacy of phonological awareness training with first-grade students who have behavior problems and reading difficulties. *Journal of Emotional and Behavioral Disorders, 9(4),* 219-231.

*Lenz, B. K., & Hughes, C. A. (1990). A word identification strategy for adolescents with learning disabilities. *Journal of Learning Disabilities, 23(3),* 149-158.

*Levendoski, L. S., & Cartledge, G. (2000). Self-monitoring for elementary school children with serious emotional disturbances: Classroom applications for increased academic responding. *Behavioral Disorders, 25(3),* 211-224.

*Lin, W. H. (1995). The effects of reading comprehension strategy instruction on reading achievement of junior-high students with reading disabilities. *Bulletin of Special Education, 12*, 235-259. (In Chinese)

Lloyd, J. W., Bateman, D. F., Landrum, T. J., & Hallahan, D.P. (1989). Self-recording of attention versus productivity. *Journal of Applied Behavior Analysis, 22(3),* 315-323.

Lloyd, J. W., Eberhardt, M. J., and Drake, G. P. (1996). Group versus individual reinforcement contingencies within the context of group study conditions. *Journal of Applied Behavior Analysis, 29(2),* 189-200.

*Lloyd, J. W., Hallahan, D. P., Kosiewicz, M. M., & Kneedler, R. D. (1982). Reactive effects of self-assessment and self-recording on attention to task and academic productivity. *Learning Disability Quarterly, 5,* 216-226.

*Lovitt, T. C., & Hansen, C. L. (1976). The use of contingent skipping and drilling to improve oral reading and comprehension. *Journal of Learning Disabilities, 9(8),* 20-26.

*Lu, M. C. (2000). The effectiveness of instruction in stem-derivation in Chinese on elementary students with severely word-recognition difficulties. *Bulletin of Special Education, 18*, 207-235. (in Chinese)

Ma, H. H. (2006). An alternative method for quantitative synthesis of single -subject researches: Percentage of data points exceeding the median. Behavior Modification, 30 (5), 598-617.

*Maheady, L., Sacca, M. K., & Harper, G. F. (1987). Classwide student tutoring teams: The effects of peer-mediated instruction on the academic performance of secondary mainstreamed students. *The Journal of Special Education, 21(3)*, 107-121.

*Maki, H. S., Vauras, M. M. S., & Vainio, S. (2002). Reflective spelling strategies for elementary school students with severe writing difficulties: A case study. *Learning Disability Quarterly, 25,* 189-207.

*Martin, K. F., & Manno, C. (1995). Use of a check-off system to improve middle school students' story compositions. *Journal of Learning Disabilities, 28(3),* 139-149.

*McDowell, C., & Keenan, M. (2001). Developing fluency and endurance in a child diagnosed with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 34*, 345-348.

*McGinnis, J. C., Friman, P. C., & Carlyon, W. D. (1999). The effect of token rewards on "intrinsic" motivation for doing math. *Journal of Applied Behavior Analysis, 32,* 375-379.

*McWhirter, C. C., & Bloom, L. A. (1994). The effects of a student-operated business curriculum on the on-task behavior of students with behavioral disorders. *Behavioral Disorders, 19(2),* 136-141.

*Miller, D. L., & Kelley, M. L. (1994). The use of goal setting and contingency contracting for improving children's homework performance. *Journal of Applied Behavior Analysis, 27,* 73-84.

*Miller, K. A., Gunter, P. L., Venn, M. L., Hummel, J., & Wiley, L. P. (2003). Effects of curricular and materials modifications on academic performance and task engagement of three students with emotional or behavioral disorders. *Behavioral Disorders, 28(2),* 130-149.

*Miller, M., Miller, S. R., Wheeler, J., & Selinger, J. (1989). Can a single-classroom treatment approach change academic performance and behavioral characteristics in severely behaviorally disordered adolescents?: An experimental inquiry. *Behavioral Disorders, 14(4),* 215-225.

*Montague, M. (1992). The effects of cognitive and metacognitive strategy instruction on the mathematical problem solving of middle school students with learning disabilities. *Journal of Learning Disabilities, 25(4),* 230-248.

*Montague, M., & Bos, C. S. (1986). The effect of cognitive strategy training on verbal math problem solving performance of learning disabled adolescents. *Journal of Learning Disabilities, 19(1),* 26-33.

Myers, J. L. (1972). *Fundamental of experimental design*. Boston: Allyn & Bacon.

*Newby, R. F., Caldwell, J., & Recht, D. R. (1989). Improving the reading comprehension of children with dysphonetic and dyseidetic dyslexia using story grammar. *Journal of Learning Disabilities, 22(6),* 373-380.

*Noell, G. H., Witt, J. C., LaFleur, L. H., Mortenson, B. P., Ranier, D. D., & LeVelle, J. (2000). Increasing intervention implementation in general education following consultation: A comparison of two follow-up strategies. *Journal of Applied Behavior Analysis, 33*, 271-284.

*Noell, G. H., Gansle, K. A., Witt, J. C., Whitmarsh, E. L., Freeland, J. T., LaFleur, L. H., Gilbertson, D. N., & Northup, J. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis, 31*, 659-663.

*Olympia, D. E., Sheridan, S. M., Jenson, W. R., & Andrews, D. (1994). Using student-managed interventions to increase homework completion and accuracy. *Journal of Applied Behavior*

Analysis, 27,85-99.

*O'Shaughnessy, T. E., & Swanson, H. L. (2000). A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities, 33(3*), 257-277.

*Pigott, H. E., Fantuzzo, J. W., & Clement, P. W. (1986). The effects of reciprocal peer tutoring and group contingencies on the academic performance of elementary school children. *Journal Applied Behavior Analysis, 19*, 92-98.

*Prater, M. A., Joy, R., Chilman, B., Temple, J., & Miller, S. R. (1991). Self-monitoring of on-task behavior by adolescents with learning disabilities. *Learning Disability Quarterly, 14*, 164-177.

*Rivera, D., & Smith, D. D. (1988). Using a demonstration strategy to teach midschool students with learning disabilities how to compute long division. *Journal of Learning Disabilities, 21(2*), 77-81.

*Roberts, R. N., Nelson, R. O., & Olson, T. W. (1987). Self-instruction: An analysis of the differential effects of instruction and reinforcement. *Journal Applied Behavior Analysis, 20*, 235-242.

*Robinson, P. W., Newby, T. J., & Ganzell, S. L. (1981). A token system for a class of underachieving hyperactive children. *Journal of Applied Behavior Analysis, 14(3),* 307-315.

*Rooney, K., Polloway, E. A., & Hallahan, D. P. (1985). The use of self-monitoring procedures with low IQ learning disabled students. *Journal of Learning Disabilities, 18(7),* 384-389.

*Rooney, K. J., Hallahan, D. P., & Lloyd, J. W. (1984). Self-recording of attention by learning disabled students in the regular classroom. *Journal of Learning Disabilities, 17(6*), 360-364.

*Salend, S. J., & Nowak, M. R. (1988). Effects of peer-previewing on LD students' oral reading skills. *Learning Disability Quarterly, 11*, 47-53.

SAS Institute Inc. (1984). *SAS/ETS user's guide* (5th ed.). North Carolina, Cary: SAS Institute Inc.

Schmidt, M., Weinstein, T., Niemic, R., & Walberg, H. J. (1985-1986). Computer-assisted instruction with exceptional children. *The Journal of Special Education, 19(4),* 493-501.

*Schumaker, J. B., Hovell, M. F., & Sherman, J. A. (1977). An analysis of daily report cards and parent-managed privileges in the improvement of adolescents' classroom performance. *Journal of Applied Behavior Analysis, 10(3),* 449-464.

Scruggs, T. E., Mastropieri, M. A., Cook, S. B., & Escobar, C. (1986). Early interventions for children with conduct disorders: A quantitative synthesis of single -subject research. *Behavior Disorders*, August, 260-271.

*Stevenson, H. C., & Fantuzzo, J. W. (1984). Application of the "generalization map" to a self-control intervention with school-aged children. *Journal of Applied Behavior Analysis, 17(2),* 203-212.

*Stevenson, H. C., & Fantuzzo, J. W. (1986). The generality and social validity of a competency-based self-control training intervention for underachieving students. *Journal Applied Behavior Analysis, 19,* 269-276.

*Swanson, L. (1981). Modification of comprehension deficits in learning disabled children. *Learning Disability Quarterly, 4*, 189-201.

*Thorpe, H. W., Chiang, B., & Darch, C. B. (1981). Individual and group feedback systems for improving oral reading accuracy in learning disabled and regular class children. *Journal of Learning Disabilities, 14(6),* 332-334.

*Trammel, D. L., Schloss, P. J., & Alper, S. (1994). Using self-recording, evaluation, and graphing to increase completion of homework assignments. *Journal of Learning Disabilities, 27(2),* 75-81.

*Trap, J. J., Milner-Davis, P., Joseph, S., & Cooper, J. O. (1978). The effects of feedback and consequences on transitional cursive letter formation. *Journal of Applied Behavior Analysis, 11*, 381-393.

*Weinstein, G., & Cooke, N. L. (1992). The effects of two repeated reading interventions on generalization of fluency. *Learning Disability Quarterly, 15*, 21-28.

*Wolfe, J. A., Fantuzzo, J. W., & Wolfe, P. K. (1986). The effects of reciprocal peer management and group contingencies on the arithmetic proficiency of underachieving students. *Behavior Therapy, 17,* 253-265.

*Wood, D. K., Frank, A. R., & Wacker, D. P. (1998). Teaching multiplication facts to students with learning disabilities. *Journal of Applied Behavior Analysis, 31*, 323-338.

*Wood, S. J., Murdock, J. Y., & Cronin, M. E. (2002). Self-monitoring and at-risk middle school students. *Behavior Modification, 26(5),* 605-626.

*Wu, S. S. (2001). Using a comprehensive reading strategies program to improve the reading comprehension skills of students with low reading comprehension. *Journal of Special Education 15*, 177-215. (In Chinese).

Author Note:

Yu-Jing Gao, Department of Education and Human Resources Development, Hsuan Chuang University, Taiwan; Hsen-Hsing Ma, Department of Education, National Chengchi University, Taiwan. This study was conducted by the first author in partial fulfillment of the requirements for a doctoral degree under the direction of the second author.

Correspondence concerning this article should be addressed to Department of Education, National Chengchi University, Wenshan District (116), Chi-nan Road, Section 2, No. 64, Taipei, Taiwan. E-mail: gxyzwgcd@nccu.edu.tw.

Author Contact Information:

Yu-Jing Gao
Department of Education and Human Resource Development
Hsuan Chuang University, Taiwan
Address: Hsiang San District (300), Hsuan Chuang Road, No. 48
Department of Education and Human Resource Development,
Hsuan Chuan University,
Hsinchu, Taiwan
Telephone: 011886353022555 Ext. 1280 (O)
011886918534299 (Mobile)
E-mail: yjgao@hcu.edu.tw

Hsen-Hsing Ma
Department of Education
Address: Wen-shan District (116), Chi-nan Road, Section 2, No. 64 Department of Education, National Chengchi University, Taipei, Taiwan
Telephone: 011886229387191(Office No. and Fax); 011886229396704(H)
E-mail: gxyzwgcd@nccu.edu.tw