CHAPTER 3

METHOD

This investigation involved the following steps: (a) location of studies, (b) coding of selected studies, (c) computation and analysis of the effect sizes of the selected studies.

Location of Studies

An exhaustive search of relevant single-subject literature was conducted using three different methods as follows: first, a computer-assisted literature search was conducted using the databases of ERIC, EBSCOHost, ProQuest and PSYCINFO with descriptors including academic behavior, single-subject, and behavioral modification. Second, a manual search was conducted of relevant behavior analysis and special education journals including 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). Third, the lists of references in the studies found by the above mentioned methods were reviewed to discover remaining studies not yet identified.

Studies that met the following inclusion criteria were selected for the synthesis:

1. The intervention focused on promoting academic behaviors.
2. A valid and scientific single-subject research design such as reversal (withdraw) or multiple-baseline design was employed.
3. Graphic time-series data displays suitable for calculating overlaps between baseline and intervention phase were provided.

Studies which included participants with mental retardation, development disabilities or physical handicap were omitted.

Coding of Studies

Studies were summarized in a table format and coded across a variety of variables in each of the following areas:

1. Classification of academic behaviors as dependent variables (i.e., academic engagement and achievement)
2. Classification of interventions as independent variables (i.e., computer assisted instruction, cooperative learning, mastery learning, reinforcement, self-control training, situated learning, social learning, strategy instruction, and others)

3. Characteristics of participants such as sex, educational level (i.e., preschool, elementary school, middle school, high school, and college), and classification (i.e., attention deficit hyperactivity disorder, autism, behavioral disorder, emotional disturbance, learning difficulties, and normal)
4. Characteristics of interventions including type of academic subjects (i.e., educational psychology, language, mathematics, science, and social science), settings (i.e., home, institution, and school), and type of intervenors (i.e., parent, peer, researcher/ experimenter, and teacher/ tutor/ principal)
5. Type of research design (i.e., reversal or multiple-baseline design)
6. Order of pairs of baseline-treatment phase
7. Conclusions of the study including the authors' judgments, discriminated as marked, moderate, or little effect, and outcome ratings, expressed as numeric 2 (effective), 1 (partially effective), or 0 (ineffective) and metric of effect size (i.e., the PND score and the PEM score)

Computation of Effect Sizes

In the current study, the PEM method is the chief method used to estimate effect sizes. The PEM score was computed by estimating the percentage of data points in a treatment phase higher or lower than the median of the previous baseline decided by a

desired orientation. For example, in the case where there is an increase in the performance of a target behavior such as participating in class discussion during the intervention phase. As illustrated in Figure 4, a line that passes through the median of the baseline phase can be drawn out horizontally (see the dotted horizontal line). As 6 out of 7 data points in the treatment phase are above the horizontal line, the PEM score is 6/7= 0.8571.

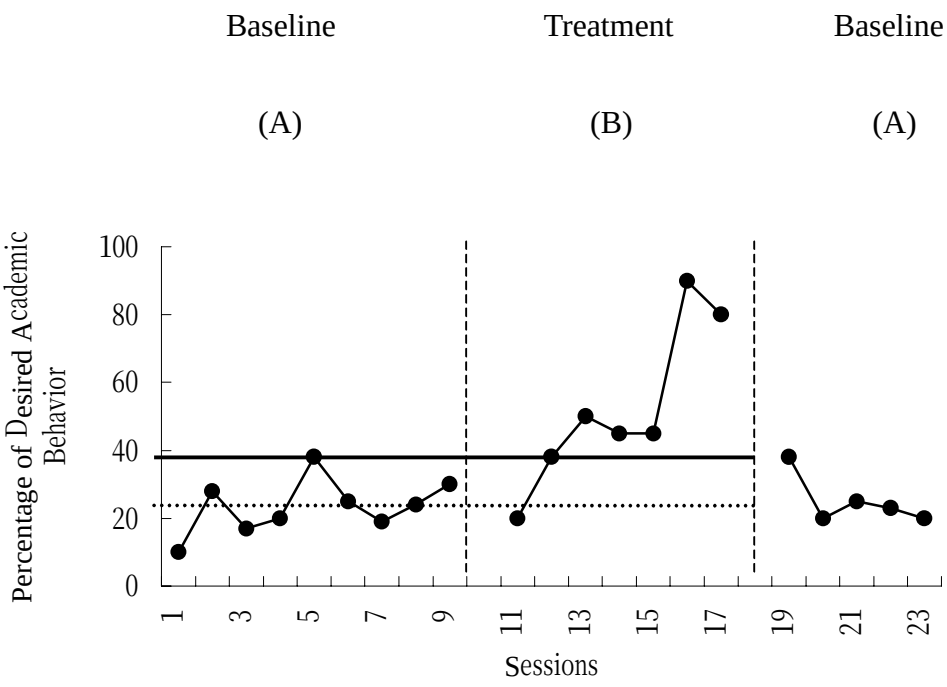

Figure 4. Computation of the PND and PEM score for a desired target behavior .

Furthermore, in the case where there is a decrease in the performance of an undesired behavior (e.g., a rate of errors) during the intervention phase, a line that

passes through the median of the baseline phase can be drawn out horizontally (see the dotted horizontal line in Figure 5). But the number of data points in the treatment phase below the horizontal line were counted. As 5 out of 7 data points in the treatment phase are below the median of the previous baseline, the PEM score is 5/7= 0.7143.

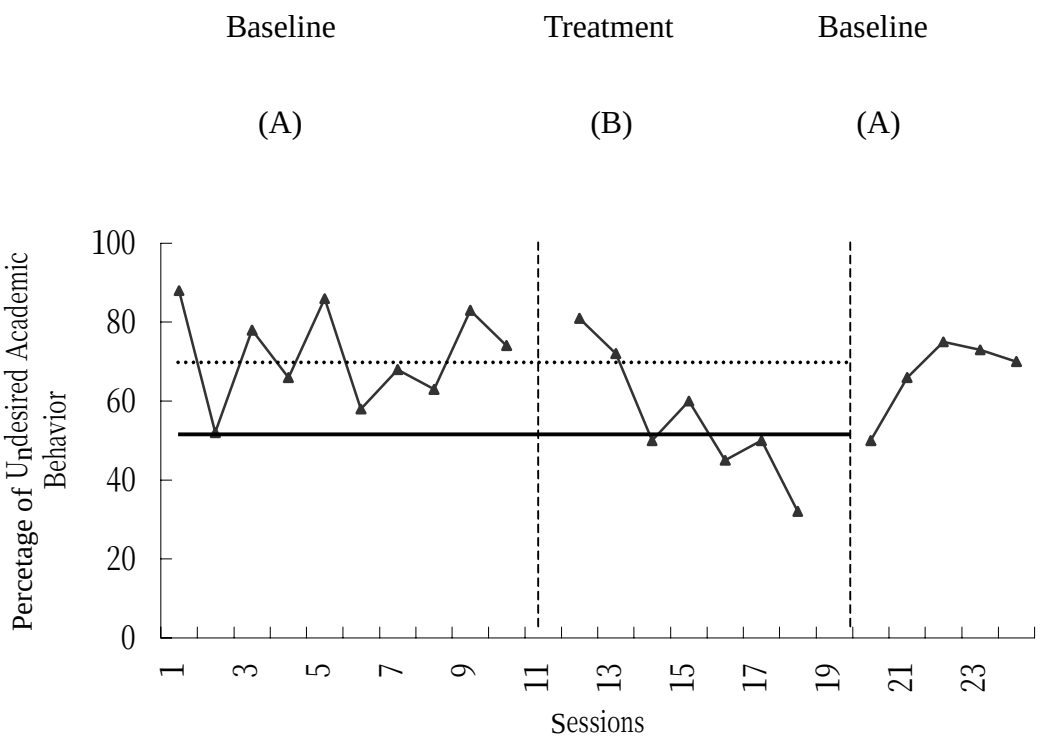

Figure 5. Computation of the PND and PEM score for an undesired target behavior.

To compare the PEM score with the PND score in this study, the PND method was also used to estimate effect sizes. A PND score is computed by calculating the

proportion of data points in a treatment phase above the highest data point or under the lowest data point in the immediately preceding baseline phase depending on an expected direction. The same examples for interpreting the PEM method (see Figure 4) are used again to explain the PND method. As demonstrated in Figure 4, there are 5 out of 7 data points in the treatment phase appear above the solid line through the highest data point in the baseline. Hence, the PND is 5/7=0.7143. And in Figure 5, 4 out of 7 data points in the treatment phase appear under the solid line through the lowest data point in the baseline. Thus, the PND is 4/7= 0.5714 while the target behavior is undesirable.

Reliability

In order to establish the reliability of the coding procedure, a doctoral student served as an independent judge to code 25 randomly selected studies. The agreement of coding between two raters (i.e., the author and the independent judge) was calculated via the following ratio procedure: the number of agreements divided by the number of agreements plus disagreements and then multiplied by 100%.

Units of Analysis

While conducting a meta-analysis, it is common to obtain multiple effect sizes from

a single study. But regarding multiple effect sizes extracted from an investigation as several independent estimates of effect may result in the overestimate of effectiveness (Faith, Allison, & Gorman, 1996). An approach proposed to deal with this issue was that of averaging total effect sizes from the same study (Faith, Allison, & Gorman, 1996).

In this study, effect sizes from the same study were averaged to resolve the issue of non-independency. And the averaged effect sizes of the identical article was regarded as an independent unit for an inferential statistical analysis such as a t-test.

In addition, effect sizes that were calculated from each pair of baseline-treatment phases were also reserved for describing effectiveness by various variables. Further, analyses of effect sizes that were calculated from each pair of baseline-treatment phases should be considered as a descriptive rather than an inferential procedure.

Validity

A comparison between the validity of the PEM score and the PND score was made by conducting the Spearman rank order correlation analysis to decide which method produced higher correspondence with original authors' judgments on intervention effects.

Detecting Orthogonal Slope Change and Outliers

The influence of orthogonal slope change in the second pair of baseline-treatment phases and outliers in the baseline were estimated by counting the percentage of their appearances.