Many social workers are anxious about the use of evaluation techniques and are particularly reticent about using statistical research. Strategies for understanding the steps involved in integrating practice and evaluation in a school social work setting, an appreciation for the need of evaluation, and knowledge of various resources on conducting evaluation and research in school social work practice are covered in this workshop document. It is suggested that school social workers implement a single-system research design into their practice and strategies for doing this are covered here. Social workers can take advantage of an abundance of existing data in a school system, such as disciplinary referrals, attendance, and grades, in setting up their system. Suggestions for interpreting data and the use of descriptive statistics are also discussed. The final step in the integrated model of practice and evaluation is reporting to the agency, the client, and the profession. It is emphasized that school social workers must share information regarding practice outcomes. It is argued that school social workers must appreciate the importance of outcome-focused research and move beyond anecdotal evidence and descriptive endeavors. Contains a list of 32 definitions. (RJM)
Many social workers are anxious about the use of evaluation techniques and are particularly reticent about the use of statistical research. In a climate of budget constraints, school social workers need to document the effectiveness of their services and programs. Workshop participants will learn how to build evaluation strategies into their own practice and thus strengthen their professional expertise and standing.

Objectives

At the end of the workshop participants will:

- Have a basic understanding of the steps involved in integrating practice and evaluation in a school social work setting.
- Have an appreciation as to the need for evaluation in advocating for and justifying school social work services.
- Have a knowledge of various resources available as to conducting evaluation and research in school social work practice.

Introduction

The push for accountability is no less for school social work than it is for other professions and social work treatment specialities. School social workers are answerable to the school district and to the students and families served by the district. There is an obligation to clearly outline school social work’s functions and methods and to provide assurances that the services are being delivered as promised (Allen-Meares, Washington, & Welsh, 1996).

Definitions of the terms used in this presentation are at the end of the main narrative.
But the task of quantitative evaluation seems daunting to the practicing school social worker. First, what is evaluation research? It is the “process of applying scientific procedures to accumulate reliable and valid evidence on the manner and extent to which specific activities produce particular effects or outcomes” (Compton & Galaway, 1979, p. 410). The use of complex, sophisticated statistical methods and procedures is not always necessary to validate practice effectiveness. The sometimes esoteric quality of these advanced statistical methods and procedures often shut out school social workers and the constituents of the school system including other school personnel, school board members, students, and parents.

A more practical approach for the school social worker is to implement and integrate a single-system research design into his or her practice. Some specific strategies to do this are detailed in the next section. Briefly, a single-system research design refers to an organized set of replicable and scientific procedures used to observe changes in an identified problem of a client or client system that is measured repeatedly over time. The given client or client system can be one or more students, a classroom, a school or an entire school system (Bloom, Fischer, & Orme, 1995).

The school social worker can take advantage of the wealth of existing data in a school system. Examples include data on disciplinary referrals, attendance, grades, handicapping conditions, achievement scores, numbers of students in special classes and remedial education, and number of students that qualify for free or reduced meals. There is a huge amount of data on budgetary and financial information in school systems and the list goes on (Weatherley, 1991).

An example of using existing school data for evaluative purposes involved a post hoc study that looked at the school records of students served by school social workers in a large school district. Regression analysis was used to compare patterns or trends in the individual student’s academic and attendance progress before and after the introduction of school social work services. The design did not allow for causal inferences, but social work intervention was seen as one of the stronger possible explanations for the improvements evinced by the students (Michals, Cournoyer, & Pinner, 1979).
Theory-based and innovative intervention approaches to combat serious problems in the schools are needed. Critical issues such as school violence, mental health and adaptive behavior issues, and school failure are serious problems that need attention and evaluation. It is essential that school social work practice be grounded to a theoretical base and not be just a collection of “hit or miss” techniques and procedures. A theoretical framework is needed for effective practice and evaluation. The practice setting of a school or school system is an ideal situation to implement and test theory-based and innovative interventions. A school setting generally provides more flexibility to assign students to different treatment and control conditions, whereas other practice environments may be more restrictive with experimental investigations (Alexander, Jr. & Curtis, 1995).

In the present climate of budget cutbacks and constraints, future educational resources may only be allocated to those services that use facts and hard data to document usefulness and effectiveness (Sabatino, Timberlake, & Hooper, 1991). And school social workers are only going to be valued by their school boards because it is believed that school social workers positively influence the educational progress of the students. More empirical evidence, such as the Michals et al. (1979) study, that documents a practical and conceptual link between academic success and adjustment and school social work services is needed.

The goals are to practice more sensitively and more effectively. The practical significance of a treatment outcome is also essential (Kazdin, 1993). As related by Bloom et al. (1995) these are the central tenets of practice—“professional action that is informed by the best available information, guided by the techniques of demonstrated effectiveness, and combined with objective evaluation components, all within the context of professional values.” (p. 15)

**Integrating Practice and Evaluation**

Effective practice is problem solving and effective evaluation is problem solving as well. Thus the integration of the two is easy and powerful. Here is how: Compton and Galaway (1979) provide a ten-step problem-solving model of practice. Bloom et al. (1995) also provide a problem-solving model of evaluation and have outlined a integrated model of practice and
evaluation using the Compton and Galaway model of practice as a guide. This integrated model also describes the steps and processes involved with single-system research designs. A precondition to actually implementing any intervention and subsequent evaluation is to establish a trusting relationship that is sufficient for practice and evaluation tasks.

The first three steps of the practice model involve identifying the presenting problem, determining the client's goals and objectives, and establishing a treatment contract. These steps dovetail into an evaluation process as well and are interrelated with assessment. Bloom et al. (1995) describe this initial three-step process in evaluative terms as conceptualizing the client's problem, specifying behavioral targets of intervention, developing a measurement and recording plan, and selecting a method or methods of measurement. During this stage of the evaluative process the target problem or objective becomes "conceptualized" and "operationalized."

The means or methods of measurement are critical parts of this process. Assessing the reliability and validity of any measure or instrumentation is crucial. The various measurements could involve behavioral observation, individualized rating scales, standardized questionnaires, client logs, and nonreactive measures. Bloom et al. (1995) is an excellent resource in outlining the various measurements available and there are other numerous resources as well (e.g., Fischer & Corcoran, 1994; Kestenbaum & Williams, 1988). Achenbach and his associates (e.g., Achenbach, 1985; Achenbach & Howell, 1993; Achenbach & McConaughy, 1987) have conducted extensive work in the area of child assessment and pathology. Two highly reliable and well-validated research and clinical tools, the Child Behavior Checklist (CBCL) and the Teacher Report Form (TRF), are extensively used by human services professionals and researchers (Achenbach, 1991).

Steps four and five of the practice model concern assessing the client's thoughts, feelings, actions, and the environmental influences affecting the client; and outlining a practice design and plan of intervention. In the evaluation process this involves initiating baseline assessment and intervention phase recording. The target or targets of intervention are measured in both the baseline and intervention phases, but during the baseline phase no intervention or treatment is
implemented, whereas during the intervention phase, one or more target-focused intervention strategies are introduced.

It is also at this point that the school social worker selects a specific evaluation design. Specific single-system evaluation designs as described by Bloom et al. (1995) can range from the simple A-B design to more complex and experimental designs. The very essence of the single-system research design is contained in these steps- the collection of repeated information on the target problems or treatment objectives. This repeated measurement points to the other terms often used for single-system designs- time-series or interrupted time-series designs. The practice or intervention design that is integrated into the research design needs to be clearly described.

And there needs to be a rationale and theoretical basis for whatever practice techniques or services are offered. As Bloom et al. (1995) relate- “There should...be clear conceptual linkages between the identified targets and the specific interventions chosen to effect them.” (p. 8) This is essential to assess the theoretical significance of the results and to add to the body of knowledge in school social work practice.

The easiest single-system research design for the school social worker to implement is the A-B design. This involves more than providing a case study because baselining procedures are implemented. An A-B design does have its limitations, however, as extensively reviewed by Bloom et al. (1995). One limitation is that it does not necessarily provide conclusive evidence that a specified intervention caused a change in the target behavior from baseline to intervention phase. However, more sophisticated experimental and combined single-system designs can approximate the statistical conclusion power of classical experimental research designs (see Orme, 1991 for an excellent discussion). Also some of the limitations of the A-B design can be compensated for by using comparison subjects, multiple designs, and a combination of multiple single-system designs and group comparison designs (Allen-Meares, 1988; Whitfield, 1996).

The next four steps in the practice model involve monitoring the progress of the intervention and providing modifications if needed, analyzing the overall intervention, maintaining the attained goal, and providing appropriate termination of service along with providing follow-up services as
needed. These steps coincide with data analysis in the evaluation paradigm. The particular procedures chosen to analyze the results of the evaluation depend upon the particular research design that was implemented. The first step of analysis for all single-system designs involves visual analysis. Visual analysis is simply "the array of one set of information relative to one or more other sets of information, so that a viewer can draw a reasonable conclusion or make a reasonable hypothesis about any relationships or lack of them among these sets" (Parsonson & Baer, 1992, p. 15).

The visual analysis is usually aided by the charting or graphing of the data. Many current integrated word processing programs (e.g., Microsoft Office, 1994; WordPerfect Suite, 1996) provide the capability to produce computer-generated graphs. Mattaini (1993) provides an informative and excellent discussion of the appropriate use of graphs and charting techniques and their illustrative power.

To facilitate the interpretation of the data, various properties of the data can be presented in graphical form. A mean line can be drawn through the data that represents the average magnitude of the data. This also aids in determining the level of the data or the magnitude of the variability of the scores or data points. A change in the level of the data across the phases is termed a discontinuity. The directionality of the data can also be assessed by graphing a trend line. A regression line is one type of trend line and represents the slope of the data for a particular phase. The data points can reflect trends that are increasing, decreasing, flat or irregular in directionality (Bloom et al., 1995).

In addition to measures of central tendency such as the mean line and assessing the trend of the data, other descriptive statistics such as measures of variability and effect size can also help to supplement the visual analysis (Bloom et al., 1995). There are also various tests of statistical significance for single-system designs. One fairly new analytic approach for time-series data is called hierarchical linear modeling or HLM (Bryk & Raudenbush, 1992; Bryk, Raudenbush, Seltzer, & Congdon, 1989). HLM was recently used to study the effects of anger control training with explosive and conduct-disordered male adolescents in a public school system day treatment
program (Whitfield, 1996). It has thus far been rarely used by school social workers, but shows great promise as it provides a model for assessing an individual’s learning rate or rate of change over time. Under certain conditions, the hierarchical structure (e.g., a student in a classroom of a school that is part of a school system) and the effects of a particular evaluative or treatment concern can be examined taking into account information from the individual and from the total group under study (Bryk & Raudenbush, 1992).

Bloom et al. (1995) is an excellent resource in describing various statistical analyses available for single-system designs. Relatively easy-to-use software (i.e., Business MYSTAT; SYSTAT, 1990) to specifically analyze single-system research data is included with the purchase of the book. Another user-friendly software resource for data analysis, particularly if some of the data will be analyzed using some group comparison procedures, is the Statistical Package for the Personal Computer (SPPC; Hudson & Hudson, 1990).

**Conclusion**

The final step in the integrated model of practice and evaluation is reporting to the agency, client and profession. It is imperative that school social workers share information regarding practice outcomes. The reluctance to evaluate contributes to the nescience of influential people as to the benefits and importance of school social work services (Allen-Mearees et al., 1996). Evaluation that is integrated with practice is not that difficult to accomplish which has been a basic theme of this presentation. Evaluation is also power and single-system research design strategies are especially powerful tools for the school social worker. Relatively rudimentary research designs and limited empirical investigations have dominated the professional literature of the social work profession in general (Glisson, 1995), and this is even more evident in the school social work literature (Sabatino et al., 1991). It is time for school social workers to appreciate the importance of outcome-focused research, both personally and professionally, and move beyond simple anecdotal evidence and descriptive endeavors.
Definitions

A-B design- The basic single-system design which has a clear baseline phase and intervention phase to make an objective determination as whether or not a change has occurred in the target problem.

Baseline (baselining)- The planned, systematic collection of information regarding the target before a given intervention is begun (referred to as Phase A in single-system designs).

Behavioral observation- The observing and recording of overt (publicly observable) and covert (internal) behaviors using direct observation and/or self-monitoring.

Case study- A precursor to a more formal single-system design in which subjective conclusions are drawn from simultaneous observation and intervention.

Causality and design validity- The presence of persistent concomitant variation between the intervention and the desired change in the target events, the statistical and practical importance of this variation and the degree to which other explanations can be ruled out as to evaluation outcome.

Client logs- An organized journal, kept by the client, of events that are relevant to the target situation, focusing on the client’s perception of the events or situations.

Conceptual definition of measurement- The dictionary meaning of what is being measured, in other words, defining a concept by using one set of words to define another word or set of words.

Descriptive statistics- The organization, summarization and description of data or a series of data points (scores) with different types of measures including measures of central tendency, variability, trend, and effect size.

Effect size- A numerical value computed to compare the amount of change from one phase to another using various formulae.

Experimental designs- A set of single-system designs that add a form of experimental control to the evaluation, in which an intervention is introduced and then removed in order to study causal effects.

Hierarchical linear modeling- A design and statistical framework which analyzes hierarchical data structures (e.g., individual student performance within a specific school of a school system) taking into account information from individual variation and relevant levels of conditions using linear regression techniques and assumptions.

Intervention (treatment)- Specifically what the practitioner does in a planned way to affect or change a target problem (identified as Phase B in single-system designs).
Individualized rating scales - Scales developed by the practitioner, in consultation with the client, to measure the relevant targets of change.

Measures of central tendency - The typical values in a series of scores or data. The median (middle score) and mean (average) are measures of central tendency most applicable to single-system data.

Multiple designs - Single-system designs where there are two or more different target problems, clients, or settings on which baseline information is collected and the intervention is successively introduced.

Operational definition of measurement - The "working definition" of a study where meaning is assigned to a concept in terms of the activities or operations necessary to measure it.

Operationalization (with a target of intervention) - Stating problems or goals in specific, measurable terms.

Phase - A time period during which a particular practice activity occurs (e.g., baseline and intervention phases).

Practical significance - Evidence that there has been meaningful change in the problem for the client.

Reactivity and nonreactive measures - Changes in a person's behavior as a result of being observed or otherwise measured constitutes reactivity. Nonreactive or unobtrusive measures include those measures that can be used without the knowledge of the person being measured.

Regression line - A trend line that represents the directionality or slope of the data. A regression line represents a line that best fits or describes the data and is calculated using a statistical formula.

Reliability - A general term for the consistency of a measurement. Different reliabilities of a measure can be assessed (e.g., test-retest reliability, alternate-forms reliability, etc.).

Single-system designs - A set of empirical procedures used to observe changes in an identified target that is measured repeatedly over time. This term has also been referred to as idiographic research, N = 1 research, single-subject design, single case-study design and time-series research.

Statistical significance - Evidence that there is a systematic difference in a target under different conditions and that this difference is not likely to be attributable to haphazard or chance factors.

Standardized questionnaire - A measure that involves the administration of the same questionnaire items (or other stimuli such as statements) to different people using the same administration and scoring procedures.
Target behaviors or events- Specifically defined objects of preventive or interventive services relevant in given situations.

Tests of statistical significance- Specific statistical procedures that are used to evaluate whether there is a systematic difference in a target between conditions.

Theoretical significance- When causal predictions or behavioral patterns suggested by a particular concept or theory of practice coincide with specific events during the course of the study.

Trend- An indication that a target behavior is either increasing or decreasing over time. This is readily observed by graphing the data and summarizing the pattern with a trend line.

Validity of a measurement- An indication of whether and how well a particular measure actually measures what is intended. The validity of a measure can involve many forms including face validity, content validity, construct validity and criterion validity.

Variability- The amount of dispersion or variation within a set of data or scores. The range of the scores and the standard deviation of the scores are two measures of variability.

Visual analysis- Simply looking at the data and finding discontinuities and other patterns that aid in the interpretation of the results. This is best accomplished by placing the raw data in graph form.
References


Other Sources Consulted


