

# **Instructor's Manual and Test Bank**

*for*

## **Evaluating Practice Guidelines for the Accountable Professional**

Sixth Edition

Martin Bloom  
University of Connecticut

Joel Fisher  
University of Hawaii, Manoa

John G. Orme  
University of Tennessee, Knoxville

**Allyn & Bacon**

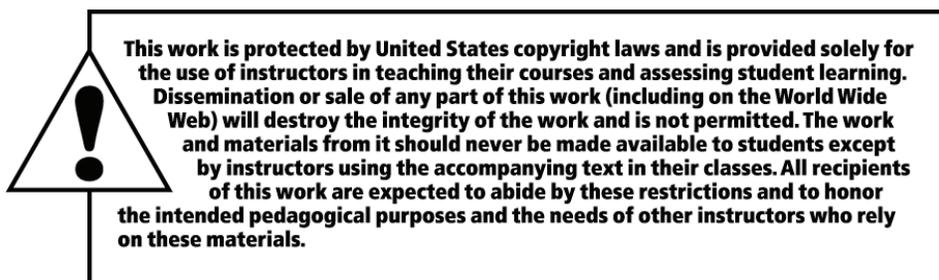
Boston New York San Francisco  
Mexico City Montreal Toronto London Madrid Munich Paris  
Hong Kong Singapore Tokyo Cape Town Sydney

Copyright © 2009 Pearson Education, Inc., publishing as Allyn & Bacon, 75 Arlington Street, Suite 300, Boston, MA 02116.

All rights reserved. Manufactured in the United States of America. The contents, or parts thereof, may be reproduced with *Evaluating Practice: Guidelines for the Accountable Professional*, Sixth Edition, by Martin Bloom, Joel Fisher, and John G. Orme, provided such reproductions bear copyright notice, but may not be reproduced in any form for any other purpose without written permission from the copyright owner.

To obtain permission(s) to use the material from this work, please submit a written request to Pearson Higher Education, Rights and Contracts Department, 501 Boylston Street, Suite 900, Boston, MA 02116 or fax your request to 617-671-3447.

10 9 8 7 6 5 4 3 2 1      12 11 10 09 08



Allyn & Bacon  
is an imprint of



[www.pearsonhighered.com](http://www.pearsonhighered.com)

ISBN-10: 0-205-61210-5  
ISBN-13: 978-0-205-61210-9

# Contents

Part I	Notes to Instructors	1
Part II	Evaluating <i>Evaluating Practice</i>	17
Chapter 1	Integrating Evaluation and Practice: Introduction to Single System Designs	18
Chapter 2	Basic Principles of Conceptualization and Measurement	23
Chapter 3	Specifying Problems and Goals: Targets of Intervention	27
Chapter 4	Developing a Measurement and Recording Plan	30
Chapter 5	Behavioral Observation	34
Chapter 6	Individualized Rating Scales	39
Chapter 7	Standardized Scales	42
Chapter 8	Logs	47
Chapter 9	Reactivity and Non-Reactive Measures	50
Chapter 10	Selecting a Measure	53
Chapter 11	Basic Principles of Single-System Designs	56
Chapter 12	Baselining: Collecting Information before Intervention	61
Chapter 13	From the Case Study to the Basic Single-System Design: A–B	65
Chapter 14	The Experimental Single-System Designs: A–B–A, A–B–A–B, B–A–B	68

Chapter 15	Multiple Designs for Single Systems: Baselines, Targets, Crossovers, and Series	72
Chapter 16	Changing Intensity Designs and Successive Intervention Designs	76
Chapter 17	Designs for Comparing Interventions	79
Chapter 18	Selecting a Design	82
Chapter 19	Basic Principles of Analysis	85
Chapter 20	Visual Analysis of Single-System Design Data	90
Chapter 21	Descriptive Statistics	94
Chapter 22	Tests of Statistical Significance for Single-System Designs	97
Chapter 23	Computer Analysis of Single-System Design Data	100
Chapter 24	Selecting a Procedure for Analyzing Data	106
Chapter 25	Not for Practitioners Alone: Evaluation for Clients, Administrators, Educators, and Students	108



# PART I

## NOTES TO INSTRUCTORS

### INTRODUCTION

This *Instructor's Manual* has a number of purposes. First, and perhaps somewhat untypically for instructor's manuals, we want to give you a little taste of our philosophy about teaching this material. You, of course, don't have to agree with that philosophy in order to make good use of this text, but we do believe in as much openness as possible, in general, with our clients, with our students, and, of course, with you, our colleagues. Second, we will provide you with an overview of the text. Third, we will introduce the resources on the **CD-ROM** that comes with this book. We have a number of exciting, useful, and to some people, possibly even entertaining resources that you can use in class to help students understand the text. Fourth, we want to introduce **trigger questions**, which are intended to peak the curiosity of the reader to a new chapter. There are (or may be) some right answers, but we don't present them. Hopefully, the chapter itself will provide the basis for the reader to discover answers for him or herself. Trigger questions speak directly to students/readers and verbalize what may be going through their mind. You might like to introduce these questions at the beginning of a class when you tackle new chapters, and let students mull over possible answers during the class period. You probably won't have to answer any of them yourself. Fifth, we will briefly discuss the format for the bank of test questions and suggested exercises for each chapter. Sixth, we will discuss some ideas about grading, both in general, and for a course using the material in the text. Seventh, and finally, we will provide a very brief overview of the rest of this *Instructor's Manual*.

### A PHILOSOPHY OF TEACHING EVALUATION USING SINGLE-SYSTEM DESIGNS

Every instructor has a philosophy of teaching, a reasoned approach to the art of conveying information and methods for conceptualizing problems, evaluating actions taken toward problem solving, and reaching decisions. However, it takes time before this philosophy becomes explicit. We three, in combined time, have been teaching for about a century, with much of that devoted to specific instruction in research and evaluation. Indeed, it was based on these experiences that Martin and Joel wrote the first edition, and then invited John to join them in the second and subsequent editions. We each have our own distinct ways of teaching, and so this Instructor's Manual is blessed (or cursed) with alternative ways of achieving various goals. This means that readers should feel free to select among the suggestions for the ways that best fit their own style and philosophy of teaching and testing. However, we three agree on some basic statements of a philosophy of teaching evaluation:

- 1) In a time of managed care and corresponding pressures to perform more with fewer resources in less time, we want to convey to helping professionals the relatively simple yet effective evaluation methods that will enable them to understand the changes taking place in client systems, to document their successes (and failures-for

further examination and correction in new cases), and to build a repertoire of strong methods for helping people help themselves. We believe that **single-system designs** (SSDs) are an ideal way to achieve these goals. Hence, we wrote the operational steps for each phase of the thinking and doing of SSDs as simply as possible.

- 2) We believe in the axiom of seeing the person-in-the-(social and physical)-environments as a basic unit of analysis, and so we try to give equal concern to evaluating person-factors and environment-factors, especially the interpersonal and social environments in which much of human behavior takes place, so as to sample from the whole situation in which clients exist.
- 3) We support a variation of the strengths perspective (Cowger, 1994; Lopez & Snyder, 2003; Saleebey, 1992), which suggests the importance of locating strengths in clients rather than only weaknesses and problems, so as to gain the motivated support of the clients in their own problem solving. This is not to deny the existence of problems, nor to neglect measurement of them. But we believe that to accentuate the positive is both good practice and ethical evaluation, particularly with minority or oppressed groups. We will frequently suggest accentuating the positive in defining targets, designing arrangements in which to measure them, and in arriving at decisions such as promoting further strengths.
- 4) While we take an eclectic position on guiding approaches for practice (Fischer, 1978), we also recognize the strength of a cognitive-behavioral perspective for teaching purposes. This position argues, first, that general information needs to be conveyed to students. This would include major principles and general assumptions, as conveyed through concepts, propositions, and theories or methodologies. The three major sections of the book each have an introductory chapter providing these principles: principles of conceptualization and measurement (Chapter 2); principles of single-system design (Chapter 11); and principles of analysis (Chapter 19). Chapter 1 is an introduction to the whole text.

Second, the cognitive behavioral perspective suggests students be tested to help them determine whether they have comprehended, stored, and are able to retrieve the general information we have tried to convey. For this purpose, we have provided this Instructor's Manual with a wide variety of testing materials. These tests are not ends in themselves, but they can be used sequentially to help students learn what they know and what they haven't yet mastered. Feedback should be part of the teaching process-not merely a grade on a paper or quiz. Each test item should be designed to test a part of the whole, and some items (like essay questions or writing an evaluation proposal) are designed to pull many of the pieces of the evaluation process together. We recommend careful use of feedback and revisions of essays or proposals as needed, based on careful comments by the instructor. To make such careful comments takes time, and so we propose carefully designed brief writings and brief tests so that the instructor can make the feedback comments that will be most useful for students. Sometimes, short, frequent tests can motivate students to keep up with the reading and to gain competency over small portions of the text. This mastery and progress should help in their sense of self-efficacy with regard to evaluating practice, our overall goal. With tests, the instructor might call attention to certain

patterns of problems, if such exist. For example, relatively more items passed on the behavior measures than on the standardized measures might suggest the need for more study on the latter.

The third aspect of the cognitive-behavioral perspective of teaching is critical; this concerns the individualization or personalization of the knowledge. This means to ask students how they are specifically going to use the information in their own practice. What will an A-B design look like in their hands? What targets, what operational measures, what graphs, and so on? This involves ample practice to translate the abstract discussion of the text and lectures into specific instances and examples from the students' own practices. How would they use this concept or method or design in their own practice situation (or, where students are not in practice, we might ask them to evaluate constructive changes they might suggest in their own lives, etc.)? Some of the best class discussions come from student examples of measurement or design or analysis problems with their on-going cases. Encourage this group problem solving, as a way to teach students to think through complex issues. It also takes the pressure off the instructor who can never be all-knowing, in any case.

The fourth aspect of our perspective on teaching involves reinforcement for the individualized application of the general ideas to a specific case or group situations. For the instructor, this involves getting "inside" the students' projects and understanding – perhaps more fully than the students themselves understand – what the projects are all about, and how best they might use evaluation to reach those goals. Again, we emphasize accentuating the positive in what the students have constructed, and phrasing our suggested corrections in the form of "have you considered doing X by Y means as a way to achieve Z goal?" And "this part of the proposal seems very workable, but I think you will need to reconsider that part..."

These four aspects are relatively easily adapted to the teaching situation, with the result that students become more involved in their own learning.

- 5) We considered long and hard about the ordering of the chapters -as packages of materials to be learned. We would have liked to teach everything all at once, but in fact we have to distribute the learning materials in some linear order so that students can absorb component pieces. We also have to attempt to put the pieces back together again, and so at the end of each Part, we have a chapter on selection: selecting measures (Chapter 10), selecting designs (Chapter 18), and selecting procedures for analysis (Chapter 24). But we also suggest modeling holistic evaluation, sharing with students what you as the instructor have done, or have read about, as an integration of the various pieces and steps of the evaluation process, in a way that might be similar to what we tried to accomplish in the *Prologue*, which is now included on the CD portion of this textbook. The overall goal is to have students see evaluation as a single cloth of which there are various threads-all of them hanging together. By emphasizing the primary colors, we simplify this process of integration; by teaching specific details, we enable students to add greater depth and variety to their methods of evaluation of their own practice.

## OVERVIEW OF THE TEXT

It is useful for the instructor to have a sense of the whole text in order to present the various parts in a meaningful way. You have, no doubt, read or scanned the text, but it may be helpful if we shared with you our plans for organizing the materials as we did. You may find it more helpful to package the learning process somewhat differently, which is fine. There is no one right way to teach. But we describe here (and in the Table of Contents of the book) what parts we have combined in what order. Sometimes, specific details may not be located in places where you think they belong. So feel free to move them (or have students read special pages in other chapters) to put whatever pieces you think belong together in the same teaching unit.

We have provided a flow diagram in the inside front cover of the book, and then again, near the end of Chapter 1, where we also give a rationale for the placement of specific chapters. *Evaluating Practice* truly is focused on a problem-solving approach in which general evaluation methods are creatively applied to particular situations, both individual cases and collective units like families, neighborhoods, and organizations. Therefore, whether the student works with individuals face-to-face, or in groups, or with families, or with whole neighborhoods or organizations, we want students to be able to take the abstract ideas and principles of single-system design as one form of evaluation, and apply them to a wide variety of contexts. We have tried to show in the *Prologue* on the CD the connection of the problem solving steps of practice with the problem-solving steps in conducting an evaluation, in order to get the feeling for the parallels in thinking and actions on the part of the helping professional. As students become more facile with practice (and evaluation), we believe these parallel processes will become second nature: It will become “natural” to think specifically and clearly about what it is that we are doing with and for the client. In what specific ways will I know how these target events are changing-including the value judgments about what is the right direction and amount of change? How do I know when I have reached closure of attaining the targeted objective-and equally important these days? How can I document these changes for people in managed care offices who may not be helping professionals themselves?

**Exhibit P1**, in the Prologue located on the CD attached to this book, shows the connectedness of practice, ethics and evaluation, and how the text has been constructed to emphasize the parallel steps in problem solving that go on in all three. It is often helpful to begin a class with a client or a group situation presented by a student. The instructor can literally go down Exhibit P1 and ask parallel sets of questions, given the establishment of a basic rapport between client(s) and worker, or between the group (such as a family) and the worker. What were the presenting problems? How did you conceptualize these problems and what ways can you see to go about measuring their existence, level, and change?

What is the client’s (or clients’) goals and objectives in this case (or situation)? How did the student-practitioner specify these goals and objectives, that is, how did the student turn a general wish by a client (or a vague discontent by the family) into a concrete target to be addressed through their mutual efforts?

Did the student-practitioner develop a contract with the client (or clients) on who was to do what, with whom, and under what conditions? Depending on the terms of this contract, the instructor can ask who was to observe (measure, count) the kinds of behaviors, feelings, or thoughts of the clients? How were these observations to be conducted (with indications of client-observation, practitioner-observation, standardized instruments, client logs, and such)?

Another major practice step is to ask what the clients thought, felt, or did with regard to the (social) environment that was a part of the target situation. Or, ask how the group relates to its several social and cultural environments. This general practice question asks about the whole ecology of circumstances, the sum of which is pushing or pulling the clients (or group) in untoward directions. This ecological picture is needed in order to design a set of circumstances that controls some of these ingredients; this is the design of the study. By controlling some factors, and introducing others at a later time, it may be possible to tease out causal factors that need to be changed in order to improve the situation for the clients.

The student-practitioner might be asked how he or she monitored changes in his or her clients or the group as a whole, and whether these changes were in the positive or negative direction. Were particular objectives attained as milestones to the overall goal? At what point did the student-practitioner decide to change or to terminate the intervention plan? On what basis? This basis is some form of analysis of the available data, as compared to identified goals and objectives.

What sort of follow-up is needed (and possible)? Can the clients (or family) take from this service experience enough knowledge, skills, and motivation to deal successfully with future challenges?

What kind of report is the student-practitioner required to present at the end of this service period? Is evidence of the effectiveness of the intervention required?

In each of the above paragraphs, we have moved from practice questions to evaluation questions, recognizing at times that they effectively were indistinguishable. Indeed, in good practice and in good evaluation, there will be a thorough blending of these two ways of helping people help themselves.

## CD-ROM RESOURCES

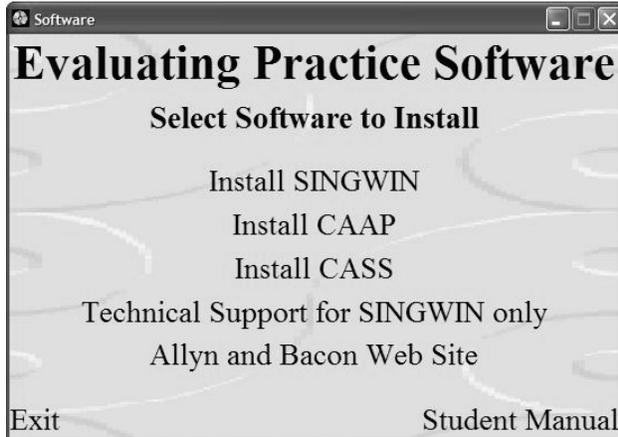
This CD is loaded with a number of resources and activities. It is our hope that these resources will make the process of learning about the topics in this book easier, more interesting and, dare we say it, fun.

The resources include the computer programs **CASS**, **CAAP**, and **SINGWIN** that you'll be reading about and using throughout this edition of *Evaluating Practice*.

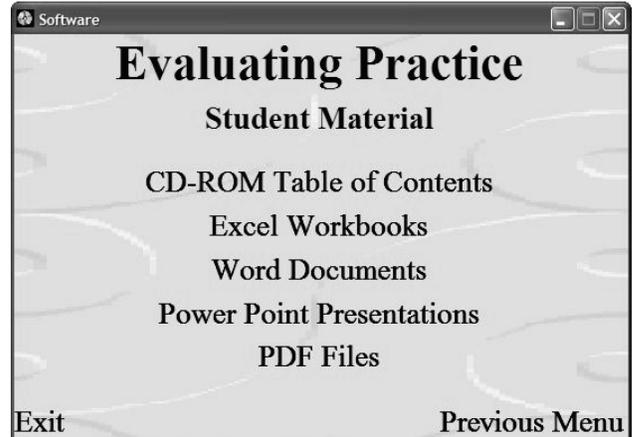
In addition to **SINGWIN**, **CASS**, and **CAAP**, this CD contains a number of other materials that we hope you'll find useful in learning and using single-system designs to evaluate your practice. These include Microsoft Word documents, Excel workbooks, Power Point presentations, and Adobe PDF files, and this document describes these materials.

### ACCESSING RESOURCES ON THE CD

1. Insert your CD-ROM and you'll see the following:



2. Click "Student Manual" and you'll see the following:



3. Notice the five menu items, each of which corresponds to a folder containing files:

#### **CD-ROM Table of Contents**

#### **Excel Workbooks**

#### **Word Documents**

#### **Power Point Presentations**

#### **PDF Files**

We describe the contents of these folders in more detail below. Each of these folders contains multiple files. To find the files, simply click on the appropriate menu item, then click on the file in which you're interested. Click on it to review the contents.

Also, notice the Word document, "CD-ROM Table of Contents"; this is the document you're reading now.

The Word documents, Excel workbooks, and Power Point presentations were developed and tested on Microsoft Office 2003. They probably work on other versions of Microsoft Office, but they haven't been tested on other versions. We apologize in advance if they don't. Please email John Orme to report any problems ([jorme@utk.edu](mailto:jorme@utk.edu)).

Now, let's turn to a description of the files contained in each of the four folders.

## WORD DOCUMENTS

---

All of the Word documents described in this section are contained in the "Word Documents" folder. Just click on this menu item and you'll see a list of the Word files described in this section. Then, just click on the file you want to open. The file contains instructions for its use. (You need to have Microsoft Office installed on your computer to open these documents.)

- **Behavioral Observation Forms**

This is a Word document that contains copies of the behavioral observation forms described and illustrated in Chapter 5. Feel free to use and modify as needed.

- **Bibliography**

This is a Word document that contains a bibliography of published single-system designs. It's not exhaustive by any means, but we hope that it gives you a place to start reading some of the interesting ways in which these have been used.

- **Client Consent Form**

This is a Word document that contains the client consent form discussed in Chapter 25.

- **Goal Attainment Scaling Form (GAS)**

This is a Word document that contains a blank Goal Attainment Scaling Form (GAS). Feel free to use it and modify as needed.

- **Interactive Exercises**

This is a Word document that contains interactive exercises designed to dynamically illustrate the relationship between single-system design data, statistics computed from these data, the interrelationships among different statistics, and the graphical representation of these data. This file contains instructions for its use. When you open this file, you can scroll to the table of contents contained in this file, click on a section, and go right to that section.

These exercises are dynamic in the sense that you can change the SSD data in this document, and when you change it, related numbers and associated charts will be updated and changed automatically. In each example, data, statistics, and charts are together on one screen. So, you can examine a wide variety of *what if* scenarios and immediately see what happens.

These exercises were created by first constructing the Excel workbooks described below, and then embedding these workbooks in this Word document. However, you don't have to know anything at all about Excel to do these exercises, and you don't need to use these workbooks unless you want to look at them and see how they're constructed.

- **Intervention Plan**

This is a Word document that contains a form that will help you practice what we preach. It was developed to parallel the suggested activities in the book, starting with Chapter 3 and the discussion of identifying objectives and goals, through developing a measurement

plan and measures of your target to developing interventions and techniques, an evaluation plan and, finally, your analysis of your results. We urge you to use this Intervention Plan with a real or even hypothetical client or on a self-change program as you read through the book. This will provide you with the real life learning that is such an essential element of comprehensive learning of any topic. Simply fill in the places on the Intervention Plan that correspond to where you are in moving through the book.

The Intervention Plan is designed to be used on your computer in Word. Simply click on “**File**” then “**Save As**” and you can save it in Word as “**Intervention Plan**” or any other designation that would remind you where to find it among your Word files. Then, you can just keep a running record of everything you do with your “client.” (Remember that the bottom “Drawing” toolbar in Word allows you to draw lines and other figures in all possible directions, so that you can create the phases of your single-system design, described in the next chapter. Just right-click on the toolbar at the top and then click on “Drawing”).

Another way to use the Intervention Plan is simply to copy it and have it available as a regular paper form. If you use it this way, we’d suggest that, before you copy it, you use your “enter” button to provide more space between categories. That is because when the form is used on your computer, Word automatically creates more space when you need it.

- **Statistical Computations**

This is a Word document that shows you how to calculate the statistics discussed in the 6<sup>th</sup> edition of *Evaluating Practice*. When you open this file you can scroll to the table of contents contained in this file, click on a section, and go right to that section.

- **Statistical Process Control Charts (SPC)**

This is a Word document that contains information about Statistical Process Control (SPC) charts beyond that discussed in Chapter 22 of the 6<sup>th</sup> edition of *Evaluating Practice*.

- **SINGWIN Frequently Asked Questions**

This is a Word document that contains questions we have been asked about SINGWIN, and our answers to these questions.

- **Walmyr Assessment Scales**

This is a Word document that contains 18 of the Walmyr Assessment Scales (WAS) discussed in Chapter 7 of the 6<sup>th</sup> edition of *Evaluating Practice* and in previous editions. When you open this file you can scroll to the table of contents contained in this file, click on the name of a measure, and go right to that measure. Note that the Walmyr Assessment Scales are copyrighted and should not be used without purchasing them or otherwise obtaining permission to use them. These scales are reproduced for illustrative purposes only. If you want to use these scales you can order copies for a very nominal fee from <http://www.walmyr.com/>.

- **WALMYR RC**

This is a Word document containing a table listing coefficient alpha, clinical cutoff, and standard error of measurement (SEM) for many of the WALMYR scales (Hudson, 1997, p. 7). Also, for each of these scales it reports Reliable Change (RC), which is discussed in Chapter 7. RC indicates minimum change indicating reliable change (RC). So, for example, for the Generalized Contentment Scale a change of 12 points would indicate reliable change, but a change of 10 points would not.

- **Orme's Course Syllabus**

This is a Word document that contains John Orme's practice evaluation course syllabus. Feel free to use and modify it for educational purposes.

- **Fischer's Course Syllabus**

This is a Word document that contains Joel Fischer's practice evaluation course syllabus. Feel free to use and modify it for educational purposes.

---

## EXCEL WORKBOOKS

All of the Excel workbooks described in this section are contained in the "Excel Workbooks" folder. Just click on this menu item and you'll see a list of the Excel files described in this section. Then, just click on the file you want to open and, where needed, the file contains instructions for its use. (You need to have Microsoft Office installed on your computer to open these workbooks.)

- **Running Case**

This is an Excel workbook that was developed to dynamically illustrate contextualized, response-guided, single-system design practice using a running case. The name of this workbook is *Running Case.xls*. This file contains instructions for its use.

- **Excel Scoring Program for Walmyr Assessment Scales**

This is an Excel workbook designed to compute total scores for the Walmyr Assessment Scales. The name of this workbook is *Walmyr Scoring.xls*. After specifying items that should be reverse scored enter data from any of the Walmyr Assessment Scales and the total score will be computed automatically.

- **Excel Scoring Program for the Center for Epidemiological Studies Depression Scale (CES-D)**

This is an Excel workbook that contains the CES-D scale and it can be used to compute a total score for the CES-D. The name of this workbook is *CES-D.xls*.

- **Visual Analysis**

This is an Excel workbook that was developed to dynamically illustrate the visual analysis of SSD data. The name of this workbook is *Visual Analysis.xls*. This file contains instructions for its use.

- **Excel Workbooks**

This is a folder that contains multiple Excel workbooks. We designed these workbooks primarily to be used in the *Interactive Exercises* document described above. To a limited extent they can be used to chart and analyze single-system design data. For example,

most of these workbooks are limited to the situation where you have an A-B design with 15 or fewer data points total for baseline and intervention. SINGWIN is a much more integrated, comprehensive, and versatile program for analyzing single-system data, but you might find these workbooks useful under some circumstances. Also, we realize that many readers of this book use Excel, and these workbooks provide examples of how Excel can be used to chart and analyze single-system design data.

The *Excel Workbooks* folder contains the following workbooks, the workbook names correspond to statistics and charts discussed in Part IV of this book, and each workbook contains instructions for its use:

- |                       |   |
|-----------------------|---|
| ➤ autocorrelation.xls | ➤ moving average.xls                                    |
| ➤ binomial.xls        | ➤ p-chart.xls   |
| ➤ boxplot.xls         | ➤ Running Case.xls                                      |
| ➤ c-chart.xls         | ➤ t-test.xls  |
| ➤ CES-D.xls           | ➤ Visual Analysis.xls                                   |
| ➤ cdc.xls             | ➤ Walmyr Scoring.xls                                    |
| ➤ chi-square.xls      | ➤ X-Bar-R-chart.xls ( <i>also creates the R-chart</i> ) |
| ➤ descriptives.xls    | ➤ X-mR-chart.xls  |

---

## PDF FILES

All of the Adobe PDF files described in this section are contained in the “PDF files” folder. Just click on this menu item and you’ll see a list of the PDF files described in this section. To open a file, just click on it. (You need to have Adobe reader installed on your computer to open the PDF files.)

- **Prologue from Evaluating Practice (5<sup>th</sup> ed.)**

This is a single PDF file that contains the Prologue from the 5<sup>th</sup> edition of *Evaluating Practice*. This prologue presents an integrated perspective on practice methods, theories and research studies, ethics, and evaluation of your own practice. It provides a holistic experience by telling a story in which you can imagine that you’re one of the principal actors in the evaluation process that takes place over time. The story is a very detailed adventure story of three students with one client in common, as well as other clients. They have to do an evaluation of this case, and proceed by steps through all of the stages we discuss in the book. We try to present some realistic “stumbles” as well as some good recoveries to a reasonable evaluation product.

- **Probability Tables**

There are four probability tables and these are described in Chapter 19.

## POWER POINT PRESENTATIONS

---

This folder contains Power Point presentations for teaching and learning practice evaluation. These presentations correspond to Orme's course syllabi in the "Word Documents" folder. Feel free to use and modify these in any way you like for educational purposes.

Just click on the Power Point Presentations menu item and you'll see a list of presentations. To open a Power Point presentation just click on it. (You need to have Microsoft Office installed on your computer to open these presentations.)

### TRIGGER QUESTIONS: THE BASIC TEN AND A GENERATIVE MODEL OF TEST ITEMS FOR INSTRUCTORS

This Instructor's Manual will present a number of alternative approaches to instructing and testing students. In this chapter, we present three approaches. The first is the **Trigger Questions** that we described earlier; these are intended to be thought-provoking questions that instructors can ask students in class as the students read and discuss each chapter.

The second approach is termed the **Basic Ten**. These are *at least* 10 test items for each chapter that briefly summarize the essence of a given chapter. They may be used as a quick pre-test at the beginning of a class period to provide a bit more encouragement for students to read and to study the assignments in advance, since there will be a test over the content. Or, they can be used in an in-class or take-home exam. But the Basic 10 also orient the students to what is central in a chapter. What we are trying to do is to provide instructors with a beginning list of basic questions; we expect that each instructor will reconstruct these questions to fit his/her own style of teaching.

These basic questions are short, so that precious teaching time is not reduced much. The instructor might choose to use only five items and save the others for a midterm or final exam. Some of these questions also can be take-home questions before a student reads a given chapter, so that the student can take them as a pre-test for his/her own comprehension of the chapter. You may or may not use 10 actual items; in order to conserve time, it may be only a couple of items that are weighted as a total of 10 points. The instructor can use these results in any number of ways, such as incorporating them into the total grade for the course, or even not counting them at all (but requiring that students do them as a kind of self-feedback). Another important use is to begin discussions or debates about the questions themselves; they are not always "easy." Indeed, they may require that students understand the information in depth, as some of the questions dig into the heart of the matter. So, these questions may be useful points of departure for class discussions.

The third approach may be called a **Generative Model** in which there is a continuum of types of questions, some of which serve one type of student-learner better than another. This approach is more suited to major tests or exams in the class. By mixing types of questions in one exam, one can try to balance the forms of the questions so that students will be able to express what they know, using a range of types of responses. For example, some students who write well find essay questions more accessible to them, while other students may find working mathematical problems easier. In fact, life is filled with demands that we sometimes provide in essays (e.g., proposals, letters, executive summaries, etc.) while, at other times, we have to provide "facts and figures." Some schools have students more or less able in mathematics and

research methods, while others may find literary productions more useful. An exam with mixed types of questions may be advantageous for the diversity of students we teach.

The simplest types of questions are recognition types in which all of the test items and answers are visually present. For example, one could take a list of key concepts and methods, and provide a list of definitions of these (and extra) terms, and ask students to connect term with definitions. Instead of definitions, instructors can list theories of which the given term is one instance.

**TRUE/FALSE [T/F] QUESTIONS** are a bit harder, since all of the information is given, and one has to decide whether this statement is true or false. This assumes instructors can write questions that make the answer clear, but do not give away the answer at the same time. This is hard to do. T/F questions tend to be either too easy or impossibly complex and hence ambiguous as to what is the right answer. Instructors would be wise to do a formal item analysis or, informally, to note the general pattern of correct and incorrect answers to see if the question itself is the problem, as contrasted with the content of the question. If the T/F question closely follows the text, then the student may just be memorizing answers rather than thinking them through, which requires a more complex form of question.

**MULTIPLE-CHOICE QUESTIONS** are a step harder, as there are several somewhat close possibilities to the target term. Each erroneous alternative should be selected to represent a body of knowledge so that the instructor can give feedback as to the type of knowledge problem the student has. For example, if the definition of “reliability” was requested from a set of definitions including “validity,” “utility,” and “reliability,” and if the student incorrectly circled the “utility” answer, then the instructor could indicate something of the kind of problem the student was having.

More complicated are **DEFINITIONAL QUESTIONS**. From a list of key concepts—which itself may be generated by the students as the most important terms they have encountered in the reading/lectures—the instructor selects a very few as indicators of the rest (and of domains of knowledge; for example, “reliability” would be a term from the reliability/validity domain, on the assumption that the student did not know what major term from this domain would be asked so the student would presumably study all of the major terms in that domain). This sampling of domains essentially is what we do as we test students over a body of knowledge. In writing their own definitions, the students have to supply the entire vocabulary. It would be preferable if students did not memorize definitions per se, but rather were able to put these definitions into their own words. This makes for a clearer understanding as students struggle with our jargon in order to come up with their own personalized understanding.

Another more complicated form of question is when we combine two or more of the key terms students have identified as important to them to learn and use in practice. Clearly, we are asking students to form propositions (such as hypotheses, axioms, etc.), which may be the point of the question, rather than the content of the items per se. So, instructors might ask students to form a hypothesis using two major concepts from a given list; evaluation would be based on whether a testable hypothesis has been constructed, more or less apart from the content of the terms themselves.

Sometimes instructors ask for a still more difficult response. A concept is provided (or may be chosen from the lists of key terms over which the student claims mastery) and the instructor asks the student to provide some other body of knowledge associated with that term. For example, what is the current state of research with regard to autocorrelation? These more

complex forms of statements require complex answers, and a move to short essay-type answers. This requires the student to connect some stored body of knowledge with the question.

More complicated still is the question that asks the student to combine two remotely associated bodies of knowledge in some socially useful way (which is a behavioral definition of creativity and a *very* demanding kind of question). For example, if students were asked to think of the connections between validity and reactivity, they might come up with a variety of possible linkages. Some might be off the wall wrong, while others may be more or less creative answers to a challenging question. [One kind of answer for the validity and reactivity question is that reactivity may spoil or change the meaning of validity...]

In our opinions, the highest form of question, and the one most useful for students in the helping professions, is the one that asks students to apply some body of knowledge to a concrete situation—a simulation of the reality they will soon face. A complex question will no doubt get a complex answer, parts of which may be accurate, parts not so much. Thus, the instructor has to determine in advance how to deal with a complex whole in terms of a grade. If there are specific points the instructor wants to see in any good answer, then the instructor should list these in advance of the test, and use this list to grade each specific essay answer.

A generative testing model has another meaning regarding the generative part. It also means that students can generate questions for themselves, and get practice in manipulating in imagination a large list of terms, all of which the instructor or the student has defined as important to him or her. So, for example, we sometimes ask students to provide us with, say, 50 to 100 terms that the students think are important to the portion of evaluation they are currently studying over several chapters, and which they take responsibility for understanding. Then, we collect their papers with the list of 50 to 100 terms a few days in advance of the exam, and we circle a very small number—sampling different domains as possible—of terms we think are most important among the students' lists. We give back these papers at the time of the exam. In this way, both students and teacher take part in developing the exam. Students feel much better, more in control of the stressful situation, and may learn more since they have a clear road map of what will be on the test—all of which they have defined as important to them. Then, as part of the actual exam, students write definitions in their own terms, which are relatively easy to grade. If students aren't selecting major terms— it is hard not to select major terms if a number like 50 to 100 terms is requested— then the instructor has another problem, to convey what is important in the content.

## **GRADING**

It is difficult to assign a letter grade (A=excellent, B=good, etc.) to a written document, although it is somewhat easier with research or evaluation activities to establish in advance the “correct” answer, when there are correct answers. More of the time, what we seek to reward with better grades is the actual thinking process, finding more suitable paths toward reasonable answers. It is a challenge to construct exams that permit us to give feedback to students about this kind of progress. Memorizing definitions or formulae are not necessarily critical to the thinking process that evaluation involves, although perhaps some memorization is necessary to have tools for thinking. So, we try to have a mix of items in tests, and then give students as much positive feedback as possible about where their strengths are, and also, helpful and constructive suggestions for increasing other areas to become strengths. Some answers are clearly wrong, or point in a less useful, less effective direction, and students need to be told this as well.

It sometimes is useful to ask students to review their own work for a segment of the course, and to give themselves a grade for their cognitive development and performance—not merely for their effort. Then, we ask for a brief (one page) rationale for this grade as a basis for a student-instructor discussion. Often, we may agree with the student on how he or she is thinking about evaluation and about overall performance, even when we may not agree on the specific grade. But it is important to understand how students evaluate their own work, since so much of their effort and understanding is invisible to us. Often, instructors can explain how they agree and disagree with students' evaluations, and in effect help students to see how the instructor's assigned grade was reached.

We would like to call instructors' attention to some basic grading systems that may be employed in classroom assignments, and at the same time, bring in some strong reality factors as well. Evaluators of grant proposals often are given general instructions on rating proposals. These include divisions of a theoretical 100 points into several broad domains, and then, sometimes, further specification of topics to be considered in each domain. This way, each proposal is rated on roughly equivalent criteria.

For example, we offer an adaptation from a training format for one kind of federal grant, but the applications can be generalized for most proposals. This training device specifies four major categories-but instructors may wish to add or subtract, as relevant to their particular instructional program. For each category, the reviewer is given a maximum number of points and a list of specific items to look at in making the evaluation. It would be possible to assign points for each specific learning objective in the course.

**I Objective and Need for Assistance (20 points)**

- A. Pinpoints relevant physical, social, economic, institutional, or other needs or problems requiring solution (that are also relevant to the grantor's Request for Proposals)
- B. States the principal and subordinate objectives of the project
- C. Provides supporting documentation, including relevant data and major references
- D. Identifies precise location of the project and area to be served by the proposed project

**II Results or Benefits Expected (20 points)**

- A. Identifies expected results and benefits and their consistency with the objectives of the proposal
- B. Indicates anticipated contribution to policy, practice, and theory and/or research
- C. Proposed project costs provided and shown to be reasonable in view of expected results

**III Approach (35 points)**

- A. Outlines a sound and workable plan of action pertaining to the scope of the project and details how each step of the proposed work will be accomplished
- B. Cites factors that might accelerate or decelerate the work, giving acceptable reasons for taking this approach as opposed to alternatives
- C. Describes and supports any unusual features of the project, such as design or technological innovations, reductions in cost or time, or extraordinary social and community involvements
- D. Provides for projections of the accomplishments to be achieved

- E. Lists the activities to be carried out in chronological order, showing a reasonable schedule of accomplishments and target dates
  - F. Identifies, to the extent applicable, the kinds of data to be collected and maintained, and discusses the criteria to be used to evaluate the results of the project (For a research or evaluation class, this would also include an expanded statement of the research or evaluation methodology, as it fits the nature and needs of the project.)
  - G. Lists each organization and key individuals who will work on the project, along with a description of the activities or nature of their contributions
- IV Staff Background and Organization's Experience (25 points)**
- A. Identifies background of project director/principal investigator and key project staff
  - B. Describes relationship between this proposed project and other work planned, anticipated, or underway (and also previously carried out) by the applicant

The same model might be usefully applied to classroom assignments, like mock evaluation proposals in which students indicate a given need or problem to be studied, the literature of what is known about this situation, a theoretical model that guides a specific practice-and-evaluation model, strengths and limitations of the design, etc.

For example, suppose a student was seeking funding to study her agency's approach to helping welfare clients learn the knowledge, skills, and motivation appropriate to succeeding in work situations. As a pilot project, she writes a proposal to a private foundation for funds to study a small group of women going through the agency's training program. She summarizes the problem based on an assessment of the literature and the experiences of the agency to date. She decides to employ a cognitive-behavioral model as most relevant to the problem and the clientele. She constructs a program that attempts to move from areas of strength (survivors in an oppressive climate) to areas of new experience (becoming employees who have to fulfill the demands of the employer at some satisfactory level of performance). The agency already provides some knowledge content, some skills training, and seeks to motivate clients toward the new work role. To what degree does this program work with a new class of trainees who are to enter in two months? By design, she can collect some baseline information on the particular knowledge, skills, and motivational levels of the clients before entering the training program. And, during the 16-week course, she can collect relevant information generated from various training sessions. About mid-way in the course, students seek jobs, and the student plans to follow them as they work part-time and continue in class part-time. The student hypothesizes that present methods may aid clients to get into jobs, but she questions whether the present training is adequate to keep them there, in the face of a cultural division between the legitimate demands of an employer and the perceived unwarranted demands of a former welfare client who feels abused by those demands. She wants to select a random subset of these clients, and present a psychosocial training module that addresses these cultural issues, based on the hypothesis that such additional training will keep workers at the jobs more than those trainees without such additional support. She proposes an A-B-C design on top of a multiple baseline design as her choice to demonstrate statistically significant changes for the C intervention, and sketches the graphs to be used in collecting data.

The weighting of such a proposal can be according to the instructor's preferences. For example, problem identification = 10%; literature review = 20%; theoretical framework guiding