

**PROGRAM EVALUATION**  
*A DO-IT-YOURSELF MANUAL FOR  
SUBSTANCE ABUSE PROGRAMS*

Brenda L. Wolfe, Ph.D.  
and  
William R. Miller, Ph.D.

Department of Psychology  
University of New Mexico  
Albuquerque, NM 87131-1161



This page intentionally left blank.

## **PREFACE**

This do-it-yourself manual was written with two purposes in mind. One is to show you how program evaluation can help you achieve the goals of your program. The other objective is to make it as easy as possible for you to conduct effective program evaluation in your own practical setting.

There are a number of reasons why program evaluation can be important and we will discuss them in detail in Chapter 2. For now, we would mention only one reason that is so compelling it is worth mentioning twice. Well-planned program evaluation helps you continue improving and strengthening your services and, consequently, your clients' experiences with the program. Systematic program evaluation can help you answer questions such as:

- ☞ What outreach strategies help us reach the largest population?
- ☞ Which clients are most likely to benefit from our current services, and how can the others be helped to do so?
- ☞ How can we reduce the number of clients lost during the intake process?
- ☞ How can we increase the number of people who participate in and complete our program?
- ☞ Is every step of the program necessary, or are there some time-consuming procedures that might be deleted without detracting from our effectiveness?
- ☞ How can we improve relapse prevention after people complete our program?
- ☞ Are there new developments in the field that could significantly enhance our service? Would adding one or more of these increase our effectiveness?
- ☞ Are our community education messages reaching the intended audiences?

No matter how structured, all programs continually evolve over time. They change, whether subtly or overtly, with changes in staff and client populations, in response to political and fiscal pressures, and as a result of new information and problems. Some of these changes are intentionally planned and implemented (for example, in response to new research findings) and others just seem to *grow*, nurtured by forces that are neither organized nor scientific.

Because change is continuous, it is helpful to think of program evaluation not as a one-time event but rather as an ongoing process. Continuous program evaluation helps to ensure that your program, as it evolves, also grows in its positive impact on those

you serve. Working without this kind of systematic ongoing feedback is a little like golfing in the dark. You can hit the ball but you have no idea where it lands. Similarly, you can implement procedures and see their immediate impact, but you cannot tell what happens in the long run, or whether there are consistent effects across clients. Program evaluation research provides you with a longer-term, multi-client perspective that helps remove the blindfold of working without reliable feedback. It gives you tools to ensure that your services continue to improve.

Although few professionals in the field would argue against its value, there is a common belief that evaluation research is so complex and time-consuming that it lies beyond the resources available to already overworked and understaffed facilities. This is simply not true. In fact, even a little systematic effort can go a long way, and *program evaluation need not be complicated*. We hope to convince you, as you work through this manual, that you can implement valuable evaluation research that is neither complex nor time-consuming. The many benefits of program evaluation do lie within reach of most practitioners and program administrators, though they may not realize it.

## **ABOUT THIS MANUAL**

In 1994, substance abuse intervention providers from across the state of New Mexico were invited to attend a *Do-It-Yourself Program Evaluation Workshop* offered by faculty of the Center on Alcoholism, Substance Abuse, and Addictions (CASAA) of the University of New Mexico. Practitioners and administrators from more than 40 facilities participated in the workshop. One of the needs expressed by these participants was for an easy-to-use handbook they could have as a guide to address evaluation questions on their own in different settings.

In response to that request, this manual was written to lay out the essential "how to" of program evaluation research. It focuses on the nuts and bolts of putting together sound projects to help you find the answers you need in order to continue improving your services. Our hope is to give you relatively easy, affordable research tools to serve a variety of needs you may have. We have tried to present the ideas broadly, so that they can be applied to a range of treatment and prevention programs. Our primary focus, however, is on the needs of practitioners and programs addressing substance abuse. Throughout the manual we will use the term "clients" to refer to those who are served by such programs, and terms like "practitioners" to describe those who deliver services.

We have tried to lay out the essentials of program evaluation research in an easy-to-read and easy-to-use manner. We have included enough background on each step to allow you to see its

uses beyond the specific examples we provide, but not so much that you become bogged down in theory or detail. We have inserted boxes to outline key points and define important terms.

Throughout the manual, we have also incorporated removable worksheets to help you apply the information to your own particular needs. We strongly encourage you to invest the time needed to complete these worksheets and methodically develop your ideas. Careful thought at the beginning of an evaluation project not only helps you have usable information at the end, but also makes the process smoother and more enjoyable.

One final word before we begin. You will notice two evaluation forms at the end of this manual. Just as we will encourage you to consider follow-up evaluation an essential part of *your* program, we also consider it an important part of our work. The first questionnaire asks for your reaction to reading this manual—how easy was it to read, did it address your needs and questions, and so on. With the second questionnaire, we are interested in hearing about your experiences in applying the information we've provided. We will appreciate it if you are willing to complete and return one or both of these—the first when you have finished reading this manual and the second when you have tried out your first program evaluation project with the help of this manual. Your feedback from both questionnaires will guide us in improving our own services, and for that we say thank you.

On that note, let's begin!

This page intentionally left blank.

## TABLE OF CONTENTS

<b>1—WHAT IS PROGRAM EVALUATION RESEARCH?.....</b>	<b>11</b>
DIFFERENT RESEARCH FOR DIFFERENT PURPOSES .....	12
<i>Summative Evaluation</i> .....	13
<i>Formative Evaluation</i> .....	13
LEVELS OF RESEARCH.....	14
<i>Level 1—Volume</i> .....	15
<i>Level 2—Reach</i> .....	15
<i>Level 3—Effect</i> .....	16
<i>Level 4—Value</i> .....	16
SORTING OUT YOUR OWN RESEARCH QUESTIONS .....	17
WORKSHEET #1. What would you like to know about your program?.....	20
SUMMARY .....	22
<b>2—BENEFITS, BARRIERS, AND STRATEGIES .....</b>	<b>24</b>
ROOM FOR IMPROVEMENT .....	24
WORKSHEET #2. How much do you know about your treatment program?.....	27
WORKSHEET #3. How much do you know about your prevention program?.....	27
CHOOSE YOUR QUESTIONS CAREFULLY .....	29
WORKSHEET #4. Evaluate your research ideas. ....	31
BENEFITS OF EVALUATION.....	33
WORKSHEET #5. What benefits do you expect? .....	35
BARRIERS TO EVALUATION.....	37
WORKSHEET #6. Anticipate barriers. ....	38
STRATEGIES FOR MAKING IT HAPPEN .....	41
SUMMARY .....	43
<b>3—QUESTIONS AND METHODS .....</b>	<b>45</b>
WHAT IS RESEARCH “DESIGN”? .....	45
DEFINING THE RESEARCH QUESTION .....	45
WORKSHEET #7. Define your research question.....	48
A FEW RESEARCH TERMS .....	50
<i>Dependent Variables</i> .....	50
<i>Data</i> .....	51
<i>Independent Variables</i> .....	51
<i>Predictor Variables</i> .....	53
<i>Control Group</i> .....	53
<i>Selecting and Assigning Participants</i> .....	53
THE KISS OF USABLE RESEARCH.....	55
<i>A Case Example</i> .....	56

BEGIN PLANNING YOUR PROJECT.....	57
WORKSHEET #8. Design your study.....	60
SUMMARY .....	62
<b>4—MEASUREMENT TOOLS.....</b>	<b>64</b>
COMMUNITY FEEDBACK .....	64
PROGRAM USAGE/REACH NUMBERS .....	66
CLINICAL OBSERVATION.....	67
BIOMEDICAL MARKERS OF SUBSTANCE USE .....	67
STRUCTURED INTERVIEWS & QUESTIONNAIRES .....	69
RELIABILITY .....	71
VALIDITY .....	71
<i>Face Validity</i> .....	71
<i>Convergent Validity</i> .....	72
<i>Divergent Validity</i> .....	72
<i>Predictive Validity</i> .....	72
INTERPRETABILITY .....	72
RECOMMENDED INSTRUMENTS .....	74
<i>Screening</i> .....	74
<i>Diagnosis</i> .....	75
<i>Assessment</i> .....	76
Patterns of Use.....	76
Substance Use Related Problems.....	78
Dependence Measures. ....	79
<i>Comprehensive Instruments &amp; Interviews</i> .....	80
INSTRUMENTS FOR FOLLOW-UP EVALUATION.....	81
SUMMARY .....	82
WORKSHEET #9. Select measurement instruments. ....	85
ADDITIONAL RESOURCES.....	87
<b>5—ASSESSING SPECIAL POPULATIONS .....</b>	<b>81</b>
AGE .....	81
<i>Assessment Instruments For Use With Adolescents</i> .....	81
LANGUAGE .....	83
CULTURE .....	85
GENERAL RECOMMENDATIONS .....	86
<i>Instruments Validated With Minority Populations</i> .....	88
SUMMARY .....	88
<b>6—FOLLOW-UP EVALUATION.....</b>	<b>98</b>
WHAT IS FOLLOW-UP EVALUATION? .....	98
WHY DO FOLLOW-UP? .....	99

HOW DO YOU DO FOLLOW-UP? .....	100
<i>Face-To-Face Follow-Up</i> .....	100
<i>Telephone Follow-Up</i> .....	101
<i>Mail Follow-Up</i> .....	102
WHO SHOULD DO FOLLOW-UP AND WHOM DO YOU QUESTION? .....	102
WHEN TO DO FOLLOW-UP? .....	104
A REMINDER ABOUT FOLLOW-UP MEASUREMENT INSTRUMENTS .....	105
INVOLVING CLIENTS IN FOLLOW-UP .....	106
PLAN YOUR FOLLOW-UP .....	106
WORKSHEET #10. Plan follow-up assessments .....	109
SUMMARY .....	111
<b>7—MAKING SENSE OF YOUR DATA .....</b>	<b>113</b>
PROTECT CLIENT CONFIDENTIALITY .....	113
CODE WORDY INFORMATION .....	114
PREPARE & ORGANIZE YOUR DATA .....	115
WORKSHEET #11. Decide what you will analyze. ....	120
CHECK YOUR DATA .....	122
ANALYZE YOUR DATA .....	123
<i>Describe the Dataset</i> .....	123
<i>Evaluate the Results</i> .....	125
The t-test .....	126
Analysis of Variance .....	128
Pearson Correlation .....	128
SUMMARY .....	130
<b>8—USING YOUR FINDINGS .....</b>	<b>131</b>
WHY WRITE IT UP? .....	131
WHAT NOT TO DO .....	132
WHAT TO DO AND HOW TO DO IT .....	132
THE PARTS OF A REPORT .....	134
<i>Executive Summary or Abstract</i> .....	134
<i>Introduction</i> .....	135
<i>Methods</i> .....	135
Sections Within the Methods Section .....	135
<i>Results</i> .....	136
<i>Discussion or Conclusion</i> .....	137
BUILD YOUR OWN REPORT .....	138
WORKSHEET #12. Plan your report—Introduction .....	140
WORKSHEET #13. Plan your report—Methods .....	141
WORKSHEET #14. Plan your report—Results .....	142
WORKSHEET #15. Plan your report—Discussion/Conclusion .....	143
WORKSHEET #16. Plan your report—Abstract .....	144

WHO TO SHARE IT WITH .....	146
WORKSHEET #17. Decide with whom to share your report .....	148
DIFFERENT REPORT FORMATS .....	150
SUMMARY .....	152
<b>9—WRAP-UP.....</b>	<b>139</b>
ADDITIONAL READING.....	141
<b>APPENDIX A—MEASUREMENT INSTRUMENTS.....</b>	<b>146</b>
ORDERING INFORMATION.....	158
REPRINTS OF PUBLIC DOMAIN MATERIAL .....	159
<b>APPENDIX B—ADOLESCENT &amp; SPANISH INSTRUMENTS.....</b>	<b>160</b>
ORDERING INFORMATION.....	160
REPRINTS OF PUBLIC DOMAIN MATERIAL .....	160
<b>APPENDIX C—MANUAL EVALUATIONS.....</b>	<b>165</b>
EVALUATION FORM #1 .....	167
EVALUATION FORM #2 .....	169

## 1—WHAT IS PROGRAM EVALUATION RESEARCH?

One of our main themes throughout this book is to define clearly and carefully the research questions you want to answer and the concepts you intend to measure. Thus, we will begin by defining what we mean by "program evaluation research."

**Program:** The system of intervention employed by an individual or group to stimulate positive changes in another individual or group.

When we talk about a program's *system* of intervention, we refer to all the pieces that are part of or affect your interactions with clients. This includes, but is not limited to, the treatment and prevention methods employed, materials distributed, interpersonal style of staff, administrative requirements of staff and clients, waiting lists, physical setting, length and composition of services offered, and a host of other components that affect services and their outcomes. Although some aspects of the delivery system (such as administrative paperwork required of staff) may appear unrelated to client outcome, they can nonetheless have an effect on clients via their impact on the staff's time, stress, and energy levels. Thus, the object of our discussions in this manual is more than simply what transpires between practitioners and clients. All the pieces of the delivery system interact to produce the *program* that clients experience. Thus, all aspects of the system can potentially benefit from program evaluation research.

**Evaluation research:** Methodical procedure for determining the relative impact of individual program components, component combinations, or the whole program, on the targeted individuals or groups.

Research is both *more* and *less* than most people think. Scientifically evaluating a program (or anything else for that matter) involves more than just observing it and drawing conclusions about what one sees. Evaluation research involves intentionally arranging circumstances in such a way that you can rule out or rule in reasons for what you observe.

For example, consider a treatment facility that routinely follows its clients for six months after they leave treatment. The staff believe that adding a three-month booster session will improve the six-month outcomes but the administration is reluctant to commit resources to this change without evidence that it will be of clear

benefit. A less systematic approach would be to implement a booster session with everyone and informally watch what happens. A more scientific approach might be to study the next 100 patients who come through the program and designate every other patient to receive the new booster session. Nothing else in their delivery system changes. When the 6-month follow-ups have been completed with these 100 (or 50, or even 30) clients, it is possible to compare the relapse rate for those who received the boosters with the rate for those who did not. If a lower relapse rate appears in the booster group, the staff can reasonably conclude that the booster is the reason for improvement, and rule out such things as the time of year or staff experience, since the non-booster group received the same treatment in all aspects except for the booster sessions. The benefits to clients of including the session can be weighed against the cost to the program of adding it.

In this example, a relatively simple plan allowed program staff to determine whether the addition of a booster session would be cost-effective in helping clients better maintain their treatment gains six months down the road. This brings us to the point that research is also less than most people think. Contrary to movie images of brilliant (mad) scientists working with complicated equipment and massive computers, good research is often simple. When one attempts to measure too many things in a single project, the result can be a whole lot of information—but not much that anyone can really understand or use. Looking at it this way, if you and the people you work with are *not* experienced researchers, you may very well have an advantage in that you will not be tempted to over-complicate things.

## ***DIFFERENT RESEARCH FOR DIFFERENT PURPOSES***

In pursuit of simple, understandable research results, it is important to begin every project with a clear idea of what you would like to know. We will talk at length about formulating research questions in Chapter 3, but at this point it is worthwhile to think about the general reasons why people may do program evaluation research. For the sake of simplicity, we will divide these reasons into two groups: Most projects are undertaken to determine either how well a program "works," or how it can be improved.

## Summative Evaluation

Projects that focus on outcome (how well a program "works") are called *summative* research because they ask how it all adds up (sums up) at the end. Summative evaluation is important, for example, when determining whether a new intervention produces the desired prevention or treatment effect. Summative evaluation is all about "the bottom line." The

sample summative questions on the next page illustrate the concept.

Summative research is also the type of evaluation one would look toward in deciding whether to continue implementing a particular program (is it working?) or, in the case of funding agencies, whether to fund a particular program (does the program have the impact desired, and is it worth the cost?). These types of projects would most often be carried out by people outside the organization looking in. Within a program, summative evaluation might be used to try out new prevention or treatment procedures and test their impact.

## Formative Evaluation

Not all research is conducted to determine the bottom line. Often practitioners and program administrators are interested in how their services might be improved or made more cost-effective. Sometimes there are practical decisions to be made: Should we add this component? Will our client flow be changed by a new outreach procedure? What kind of training is most important in increasing staff skills or morale? Such concerns are the focus of *formative evaluation*. This type of research focuses on the building blocks that make up a program and how each can be optimized to provide the best possible services.

### Sample Summative Questions

☞ A new anti-smoking program is developed for use in the schools. Prior to implementing it statewide, we test it in a sample of schools and compare subsequent tobacco use in those schools with a comparable set of schools that do not receive the program. If the program produces lower smoking rates in its schools, we conclude it "works."

☞ XYZ Pharmaceuticals has developed a new drug to reduce craving for alcohol. To determine whether it truly works, half of the alcoholic clients at each of three treatment centers are given the drug and the other half are not. Cravings are tracked for both groups for the duration of the study and frequencies compared at the conclusion. If the drug group reports significantly fewer cravings than the non-drug group, we conclude the drug works

The benefits to clients of formative evaluation are obvious. However, there is an equally important, but less obvious, benefit to staff. It makes your job easier. With the feedback provided by this evaluation, you can fine-tune your skills, invest your energy in the most productive practices, identify the most accurate instruments to use in your work, and generally be more productive and feel better about your performance. Examples of formative evaluation research will be given later in this manual. As you think about the evaluation needs in your own facility, consider whether you most want to know "How well does it work?" or "How can we make it work better?" If it is the former, you are looking for summative data; if the latter, your purpose is formative. In sum, summative evaluation gives you an overall outcome report and formative evaluation helps you continue to build and improve your program. In fact, both kinds of research can be important to a program and sometimes the same evaluation project can serve both purposes.

## **LEVELS OF RESEARCH**

There is one more set of definitions you may find helpful to know. These describe four levels of information gathered in evaluation research. There is a set of formal labels<sup>1</sup> that have been used for each level (effort, performance, outcome, and efficiency), but we have chosen to use names that may be easier to remember and understand: **volume** (how much service activity is generated—staff hours, patients seen, etc.), **reach** (how well does the program meet community needs), **effect** (what impact does it have on clients), and **value** (how much effect you get for the cost).

---

<sup>1</sup> James, G. (1969). Evaluation in public health practice. In H.C. Schulberg & F. Baker (Eds.), *Program evaluation in the health field*. New York: Behavioral Publications.

## Level 1— Volume

Evaluating the program at this level means looking at how much *time and effort* are expended. In other words, what volume of service activity has been generated by the program—how many staff hours are being used, how many clients reached, what populations are being served? Examples of research questions that evaluate program volume are shown here. Many programs already collect information of this kind.

### **Sample Program Volume Questions**

## Level 2— Reach

In a nutshell, this level of evaluation asks how close the volume of services delivered comes to meeting the community's needs. In other words, how many of those in need are being *reached* by services? This is the type of information that is often required to justify increased staffing or budget. Level 2 involves comparing your program volume with an estimate of need.

Volume and reach levels of evaluation provide useful information for answering questions about the "length and width" of a program. They tell you nothing,

### **Sample Program Reach Questions**

however, about what impact the program has on its target population. This next level does this.

### Level 3— Effect

Evaluation that focuses on program effectiveness tells you how well the program accomplishes its objectives in serving its population. It tells you about the program's actual impact. It is possible, for example, that a prevention might reach (be in contact with) many people, but have virtually no beneficial *effect*. The box to the right gives you a sense of the types of questions one might ask at this level.

#### **Sample Program Effect Questions**

- ☞ How many patients require no further treatment after "graduating" from the program?
- ☞ Does the introduction of an anti-drug campaign reduce the number of first-time users?
- ☞ Does attendance at AA meetings concurrent with treatment decrease relapse rates post-treatment?
- ☞ Can follow-up contacts be increased by requiring a monetary deposit at treatment outset?

It is at this level you begin to receive feedback that can improve the quality of your program and simplify your own tasks. Of course, it is always tempting to rely on "professional judgment" in judging programs. However, this is no substitute for good evaluation. Consider the following situation. A public school system implemented a drug abuse prevention program to which teachers and students alike gave enthusiastic endorsement. Students seemed to like the new curriculum and teachers were convinced it was preventing drug use. Yet, when the program was actually evaluated by comparing its students with others who did not receive the program, it was found to actually have *increased* students' drug use<sup>2</sup>. The same enthusiasm that makes program implementation a success can also obscure a program's true effects. *Effectiveness* research removes the blindfold.

### Level 4— Value

The fourth evaluation level, *value*, simply refers to the cost-effectiveness of the program (weapons manufacturers wryly refer to this as "bang for the buck"). When a high-quality product is bought at a reasonable price, one is said to have received "good value." Sample research questions about program value appear below. As you can see, they address how much impact (outcome) the program generates per unit of cost.

---

<sup>2</sup> Stuart, R. B. (1974). Teaching facts about drugs: Pushing or preventing? *Journal of Educational Psychology*, 66, 189-201.

Information from value studies can lead to effective streamlining of services as well as improvement in the number of clients served with the same budget. A good example of how value data can benefit service delivery comes from research comparing outcomes of inpatient or residential alcohol programs, which tend to be costly, with less expensive outpatient programs. The surprisingly consistent finding of

similar outcomes for residential and outpatient programs has led to greater reliance on the lower-cost options, which can make needed services available to more people within the same budget<sup>3</sup>.

Together these four levels add up to a comprehensive picture of the program. Level 1 describes the *volume* of activity and Level 2, *reach*, tells you how well that volume meets the needs of your target community. The third level, *effect*, tells you whether you have the desired impact on those being served, and Level 4, *value*, helps you evaluate whether resources are being allocated in the best possible fashion. Taken altogether, you can learn much about the workings of your program and how to strengthen it. All of these levels together, however, generate a lot of information to try to make sense of, which brings us back to our earlier point that simplicity in science can be a virtue. As you begin to think about evaluation efforts that would most benefit your organization, consider which level(s) of information are most important at this time, and focus clearly on what you need.

### **Sample Program Value Questions**

- ☞ What is the per-patient cost of Treatment X versus Treatment Y?
- ☞ How does the cost of one additional booster session compare to the cost of patients requiring repeat detoxification within 12 months post-treatment?
- ☞ Can comparable results be obtained with the low-cost drug, Makimwell, and the more expensive drug, Chargimwell?
- ☞ Will the use of paraprofessionals in the field result in the same community response at a lower cost?

## **SORTING OUT YOUR OWN RESEARCH QUESTIONS**

Once you start thinking about aspects of your program that could benefit from a better understanding of how they work and how they might be enhanced, research questions will begin to pop out at you. Exciting as this can be, there is a danger in the excitement. Good research, realistically simple research, comes

---

<sup>3</sup> Institute of Medicine, National Academy of Sciences (1990). *Broadening the base of treatment of alcohol problems*. Washington, DC: Author.

from keeping a clear picture in your mind of what information you are after and how you plan to use it. Thinking of your questions in terms of whether they need summative or formative answers, and considering what *levels* of information you need will help you remain focused and gather useful and usable information.

Take a few minutes now to start your ideas flowing and sort out your research questions by completing Worksheet #1 on page 20. Try to put at least one entry in each section of the worksheet. Even if you have no intention of doing a particular type of project, thinking through the possibilities will stimulate your ideas and sensitize you to questions that might creep into your projects that don't match your objective.

This page intentionally left blank.

WORKSHEET #1. WHAT WOULD YOU LIKE TO KNOW ABOUT YOUR PROGRAM?

LEVEL 1—QUESTIONS ABOUT VOLUME

LEVEL 2—QUESTIONS ABOUT REACH

LEVEL 3—QUESTIONS ABOUT EFFECT

LEVEL 4—QUESTIONS ABOUT VALUE

This page intentionally left blank.

## **SUMMARY**

There is more to a program than meets the eye. As all aspects of a program come together to create the program and services that clients actually experience, different dynamics come into play and things do not necessarily work exactly as you anticipated. Thus, continually evaluating the results of your efforts can be helpful in ensuring the best possible program.

As you formulate your research questions, think carefully about what you want to know and how you intend to use the information, and keep these objectives clearly in mind. Considering your questions in light of the purposes and levels of research we have discussed may help you to keep that focus.

This page intentionally left blank.

## **2—BENEFITS, BARRIERS, AND STRATEGIES**

We hope that Chapter 1 got you thinking and even a bit excited about the potential for gathering useful information about your program. Answering meaningful questions can be very stimulating and, if you take to it, one of the most satisfying tasks in your work. Even those who catch the bug and get excited about program research, however, face one of the great barriers to human change—inertia. Therapists know that getting started on any project is often the toughest step, particularly if there is no urgent reason to do so. If your program has been operating pretty much the same way for years and no one has really complained, there are going to be those who ask, "Why mess with success?" In fact, you may also find yourself wondering at times why you are going to all this trouble when nothing is really wrong with the program. There is at least one simple answer to that question: Nobody's perfect.

### **ROOM FOR IMPROVEMENT**

As long as no one is perfect, no program will be perfect. That means there will always be some room for improvement. Think about your program, your *whole* program. How many of the clients with whom you come into contact live happily ever after as a result of that contact? Can you think of areas in which you wish there were more effective ways of working? Are there administrative requirements that you would like to streamline in order to free up more time for clients? What additional information would it be helpful for you to know about your program's effect? Is what you already know based on systematically collected facts or on incidents and impressions?

Because we are human, we naturally tend to notice and remember the things that fit best with our hopes and expectations, and to overlook or forget those that make us uncomfortable. Unfortunately, this means we sometimes miss important information. The next two worksheets on pages 26 and 27, one for treatment and one for prevention programs, pose some questions for which it might be useful to know the answers. The worksheets are informal self-assessments of where you stand in having a good grasp on your program's outcomes. If someone outside the program, perhaps a funding source, asked for such information, how well could you respond? The point is that no matter how sincerely we attend to our clients, there is much we miss if we do not have a system for gathering our observations. In an era of increasing demand for accountability, programs are vulnerable

if they do not have reliable information about their volume, reach, effect, and value. Program evaluation gives you that information.

This page intentionally left blank.

WORKSHEET #2. HOW MUCH DO YOU KNOW ABOUT YOUR TREATMENT PROGRAM?

How long does the average client remain in treatment?

What percentage of clients relapse within 3–6 months of leaving treatment?

For the average client, how long is it until the first lapse?

What are the most common triggers of relapse for your client population?

What percentage of clients recycle through the program a second time?

How many clients go on to seek treatment elsewhere after leaving your program?

What are the top three reasons clients come to your program rather than others?

What are your clients' most serious concerns before and after treatment?

To what extent is your program decreasing the harm that results from substance abuse, both for your clients and for those around them?

Beyond its impact on substance use, what effects does your program have on your clients' lives more generally?

Now go back and check off the answers that you *know for a fact* (that is, for which you have hard evidence).

WORKSHEET #3. HOW MUCH DO YOU KNOW ABOUT YOUR PREVENTION PROGRAM?

What proportion of the population you'd like to reach do you actually contact?

Who is being reached, who is not, and why?

How much does contact with your program change the knowledge, attitudes, and behavior of the population you

reach?

Does your program reduce the incidence of first-time substance use, as compared with areas in which no such program is offered?

To what extent does your program contact current substance abusers and influence them to change their behavior?

Who are the people most benefited by your program?

What other beneficial changes (such as reduction in adverse health consequences) occur following implementation of the program?

Now go back and check off the answers that you *know for a fact* (that is, have hard evidence for).

Perhaps this list does not contain the questions that are of most importance to your own program. If reviewing Worksheets 2 and 3 suggests other questions you'd like to answer, so much the better. The point is that in most programs, the staff have little systematic feedback (and thus evidence) about many important issues, even those that are central to the program's mission and survival. If this is true for your agency, it does *not* mean your program isn't working. It *does* mean you are missing important information that can help reinforce your strong points and strengthen areas where you are less effective. Formative program evaluation is not about fixing a broken program. It is about improving a working program.

People are wonderful, information-processing organisms. We go through each day taking in new information, finding its fit with old information (sometimes having to ignore bits and pieces of new experiences that don't fit), and quickly drawing conclusions that allow us to move on to the next experience. This is a reasonably effective means of moving through life without getting bogged down in useless detail. We all do it. Nevertheless, we do it at the cost of missing potentially important bits and pieces that don't fit our expectations and so get overlooked. While this presents no serious problem for most of our tasks, when it comes to evaluating the effect of a program implemented by a number of people for a number of people, little bits of data missed from client to client can add up to great big missing chunks of vital information. Most of us tend to see that which is encouraging and to minimize that which is discouraging. (One of the only exceptions to this perceptual bias, in fact, is those who are clinically depressed.) This means that relying on your own and staff *impressions* that have not been *systematically* gathered can give you a somewhat inaccurate picture. This unintentional bias can rob you of important information that could improve the quality of your services and make your job easier.

## **CHOOSE YOUR QUESTIONS CAREFULLY**

At this point you probably have considered a number of questions that you believe would help your facility improve its services. Take a few minutes now to review the ideas you jotted down on Worksheet #1 and consider some of the missing information highlighted by Worksheets #2 and #3. Think about the objectives of your organization and to whom it is accountable. Consider all these things and narrow down your questions to the *two or three* that you conclude would be the most useful and acceptable to program staff. Then talk to the people who would be key participants in helping you accomplish this project. Present your thoughts to them,

focusing on both the pros and cons of each possible question, and decide as a group on the one most pressing question. (Remember, if your first project provides useful information that benefits the program, motivation will be higher for the next project.)

Once you and your group have agreed on the question that will form the basis of your program evaluation project, describe it in Worksheet #4 on page 31. As you work through the remainder of this manual, this question will be further refined and, quite possibly, modified. However, it is important that you keep clearly in mind the central information that you want from your evaluation project. If you do so, everything you do from this point on will bring you closer to the answers you need.

WORKSHEET #4. EVALUATE YOUR RESEARCH IDEAS .

REVIEW THE QUESTIONS YOU GENERATED IN WORKSHEETS #1, #2, AND #3 AND DISCUSS THE TOP TWO OR THREE IDEAS WITH YOUR COLLEAGUES. BASED ON THOSE DISCUSSIONS, SELECT THE ONE QUESTION THAT BEST ADDRESSES YOUR ORGANIZATION'S MOST PRESSING NEED. DESCRIBE THAT QUESTION HERE.

This page intentionally left blank.

## **BENEFITS OF EVALUATION**

By now you have figured out that we are die-hard advocates of program evaluation as a way to ensure your program's growth, quality, and even survival. (We will try not to repeat ourselves too many more times on that point—but we may relapse.) In addition to providing you with the information you need to continually improve your program, evaluation research can bring other benefits as well.

☞ Active involvement in research sends the message to clients that the service providers care about quality, are up-to-date in their field, and connected to important sources of information.

☞ In a way, for providers to engage in program self-evaluation is to model some of what you hope to convey to your clients. People struggle with the many challenges, pitfalls, and setbacks on the road to health. Being able to recognize and deal with shortcomings or mistakes is, in fact, a key to success in overcoming addictive behaviors. Convincing clients to accept their shortcomings and learn from them might be made easier by pointing out how the program itself is continually being evaluated and improved based on experience. Health promotion and recovery are, like program development, a continual journey that requires regular inventory-taking and course correction.

☞ An increased awareness of how various program components affect individual clients allows greater personalization of the program. It can mean the difference between trying to make all clients fit into your program, and individualizing your program to fit clients' needs.

☞ Accurate information about how the various components of the program work together to achieve your objectives often translates into better outcomes for your clients.

☞ The feedback provided to practitioners by evaluation research helps them not only improve their therapeutic skills but, equally important, provides positive reinforcement that is sorely lacking in many workplaces and workdays. Without such data, the practitioner can only guess and never really knows what she or he is *doing right*. Without feedback about success, people stop learning and growing.

☞ Program reach evaluation can provide the justification needed to increase budgets and allocate resources more effectively.

☞ The increasing accountability to which programs are held by government agencies, private boards, and third-party payers means that the demand for objective measurements will likely increase and, along with it, the competition for limited financial resources. By implementing a good evaluation process now, your facility will be better prepared for the inevitable day when you are asked for hard numbers about the success of your program.

These benefits are fairly generic and may apply, in varying degrees, to your own work setting. However, we could not possibly have addressed the most critical questions regarding your own organization because we simply do not have the in-depth familiarity that you do. Take some time to consider what you know about your program and what you would like to know about it. Consider the evaluation question you posed in Worksheet #4; how could its answer benefit your organization, your clients, and yourself? Worksheet #5 on page 35 will help you organize your thoughts. It will be helpful to have clear notes about all the ways program evaluation can help the organization when it comes to setting up your evaluation project and involving colleagues.

WORKSHEET #5. WHAT BENEFITS DO YOU EXPECT?

TO YOUR ORGANIZATION?

TO YOUR CLIENTS?

TO YOURSELF?

This page intentionally left blank.

## **BARRIERS TO EVALUATION**

Just as there are many pieces and processes that go into making up a whole program, there are many variables that enter into conducting a successful evaluation project. Although the project itself need not be complicated, the process of conducting it in your setting can become complicated by organizational, political, and human factors. Of course, not every project in every setting is complex (many are quite simple). However, you are best off giving some thought to the potential barriers before you encounter them (if you do).

While every organization has its own unique grouping of people and circumstances that may help and hinder, there are a number of common issues that tend to arise. The following table points out possible barriers to program evaluation and suggests some ways for overcoming them. As you review the table, keep in mind that each organization is unique and these barriers may or may not apply to your own setting. If they don't, you can cross them off your list. If they do, thinking them through up front will minimize their impact on your progress.

<b>BARRIERS AND POSSIBLE SOLUTIONS</b>	
<p><b>Fear.</b> For most people, the idea of <i>evaluating</i> what they do all day is a little scary. Even though we are talking about evaluating a program, there will be people in every organization who perceive the project to be a threat to them.</p>	<p>Awareness of and concern for those people who are unhappy with the project, a clear statement of the project purpose, careful planning, and good project management can effectively defuse many objections.</p> <p>In stating the project purpose, emphasize the benefits to be gained by both clients <i>and staff</i>.</p>
<p><b>Politics.</b> No matter where you work, if the organization consists of more than one person, there will be politics.</p> <p>There may be people who try to block the project to protect their position in the organization or prevent others from furthering their careers. There may also be people who try to use the project to either further their own career or sabotage someone else's.</p> <p>Questions may arise about who is evaluating whom and how the information will be used.</p>	<p>Program evaluation need not mean the measurement of success or failure of any particular practitioner's clients. <i>Program evaluation</i> means just that—assessment of how effective the <i>program</i>, or its components, may be in effecting desired change in the population it serves. It is important that this objective be clearly stated at the outset of the project and emphasized throughout its duration.</p> <p>Potential sabotage from political maneuvering can be minimized by making sure that key people in the organization are fully on board. Get them involved at the very earliest planning stages to ensure they have a sense of ownership in the project and will project that feeling to the people within their sphere of influence. It is also helpful to appoint a project manager whose duties include, among other things, keeping a sensitive ear out for potential problems.</p>
<p><b>Intimidation.</b> People believe that evaluation research requires staff who are trained in scientific research and data</p>	<p>To the contrary, useful program evaluation projects can be kept quite simple and managed by virtually anyone with reasonable organizational skills. With respect to the design of the project and</p>

<b>BARRIERS AND POSSIBLE SOLUTIONS</b>	
analysis. Few agencies have this type of staff.	interpretation of the results, this manual is designed specifically to simplify those tasks.
<b>Resources.</b> Few agencies have excess money or staff time to invest in research.	The more careful thought that goes into the design of the project, the easier it is to keep costs to a minimum. Many of the measurement instruments you might need can be obtained free of charge, research procedures can be designed to fit smoothly into current practices, and outside resources can often be obtained at little or no cost from local colleges and universities.

Of course, every workplace is as different as its collection of staff and clients. The barriers that you face may be similar to the ones we've discussed or may be completely unique to your setting. Barriers that you anticipate are always easier to negotiate than those that catch you by surprise. Think about your own organization:

- ☞ How and where are really important decisions made in your organization?
- ☞ Who are the opinion leaders, the people who have real influence?
- ☞ What standard procedures might help or hinder your project?
- ☞ How are funds and other resources allocated?
- ☞ Are there any particular individuals who might have strong objections to the project? How might they be brought on board?
- ☞ Who in the organization could be an important resource and supporter for your project?

Take a few minutes to fill in Worksheet #6 on page 38 by listing the barriers you can anticipate and solutions you believe might help. Do not be concerned if you have trouble filling in the whole worksheet. Solutions will become more apparent as you plan the details of the project itself and, we can promise, you will not feel deprived of challenges once you get going! However, the more anticipation and up-front planning you do, the smoother your road will be.

<b>WORKSHEET #6. ANTICIPATE BARRIERS.</b> <b>WHAT OBSTACLES ARE YOU LIKELY TO ENCOUNTER IN YOUR ORGANIZATION? BEGIN PLANNING SOLUTIONS.</b>	
<b>POTENTIAL BARRIERS</b>	<b>POSSIBLE SOLUTIONS</b>

WORKSHEET #6. ANTICIPATE BARRIERS.

WHAT OBSTACLES ARE YOU LIKELY TO ENCOUNTER IN YOUR ORGANIZATION? BEGIN PLANNING SOLUTIONS.

POTENTIAL BARRIERS	POSSIBLE SOLUTIONS

This page intentionally left blank.

## **STRATEGIES FOR MAKING IT HAPPEN**

In large corporations where the success or failure of a project often determines whether profit lines go up (and people get promoted) or down (and people get to job hunt), few employees pass up the opportunity to take project management courses. Between having a great idea and seeing it fulfilled, there are a lot of little steps that need to be taken. While these steps need not necessarily be difficult ones, if too many are missed, a project can be seriously compromised

Happily, an in-depth course on project management is not needed to conduct program evaluation. However, being aware of a few basics will help your project go more smoothly. (The first two points may be irrelevant to you if you work alone.)

- **Assign a project leader and project coordinator.** Any project that has more than one step to it, as all evaluation projects do, needs someone who is responsible for overseeing the bits and pieces that must combine to form a successful outcome. Just as an orchestra composed of top-rate musicians still needs a conductor to create a beautiful sound, the members of your project team need a central figure to keep them focused in a common direction and playing in time with one another. A common reason for the failure of evaluation projects is the lack of someone who is clearly responsible to get them done.

If your organization is multi-layered and there are more than just a handful of people involved in the project, it is a good idea to have both a project leader and a project coordinator. The *leader* would ideally be someone who is influential enough in the organization to help keep everybody focused on the same vision as well as have the power to reallocate resources if needed. The *project coordinator* is the person actively involved in keeping track of the bits and pieces that make up the project. He or she is responsible for making sure data are being gathered, necessary procedures and paperwork are being done, and so on. It is usually a good idea to select the coordinator from among the administrative personnel in the organization and to look for someone who has top notch organizational and *follow through* skills.

- **Get everyone on board.** Don't wait until the project plans are set in stone to get other members of the team involved. As we discussed earlier, political maneuvering and perceived threat can be significant

barriers to a smoothly running project. You can minimize these types of problems and significantly improve the quality of the project by including a variety of perspectives and areas of expertise early in the process.

- ***Plan, plan, and master plan.*** Luck may be useful if you are rolling dice in a casino, but it is close to useless in making good projects happen. Careful planning of your project from beginning to end *before you begin* is key to a smooth operation. From defining your research question to determining your materials and procedures to planning how you will interpret and distribute your findings, it is helpful to have a written master plan. This plan should be thought out in as much detail as possible, recorded, and distributed to all team members. Don't panic! This written plan need not be formal. A simple list of *to-do* items that begins with a clear statement of the project objective, followed by the *to-do's* and *who* is responsible for doing them *when* will suffice.

As you write out your plans, keep in mind that the realities of implementing them over time will often require that you adjust your strategy as you go. That's fine; but as you do, be sure to also adjust your master plan and keep everyone informed.

- ***Create a timeline and track your progress.*** Whether your project has a distinct beginning and end or is one that implements an ongoing evaluation procedure, having an agreed upon time frame for its implementation will help keep the project moving forward in a timely fashion. Including target dates for the *to-do* items on your master plan will be useful to the project coordinator as she or he reminds people of upcoming tasks, prepares reports, and so on.
- ***Assign tasks to individuals.*** When assigning the *to-do* items on your master plan, assign them to *specific people* and not to a department or "whoever has time." While the group may all be enthusiastic and eager to help, other job responsibilities will get in the way of project tasks if someone has not been designated and has accepted the responsibility. The opposite problem—too many people jumping in to do the task—will result in a duplication of effort.
- ***Schedule regular meetings and provide interim reports.*** Meetings can be one of the most useful tools in project management or one of the most deadly. Very useful are

regular, short, well-planned meetings in which the project leader or coordinator updates everyone on the project's overall status, asks for reports from team members on their areas of responsibility, and anticipates upcoming challenges as a group. Extremely unhelpful meetings are those that go on for more than an hour and during which the participants are not updated on the big picture of the project nor given the opportunity to report on their progress and, if necessary, seek guidance or assistance from other members.

## **SUMMARY**

Conducting program evaluation research does not need to be a difficult undertaking. However, it is an undertaking that requires some careful attention to detail. Investing the time to identify the specific benefits your organization and clients will receive from the project as well as the potential barriers that may spring up, will make the process go more smoothly.

If we had to summarize all the advice ever given about running successful evaluation projects, the key word would have to be **plan**. By investing the time in planning all phases of the project and keeping careful records of the plan (including how it might change over time) and tracking your progress against the plan, you will not only help the project succeed but also keep the monetary and time costs to a minimum once it gets going.

This page intentionally left blank.

### **3—QUESTIONS AND METHODS**

Here's an opening tip: As you design and implement your evaluation project, keep your research question clearly in your own mind and in the minds of everyone who is involved. A common mistake that evaluators make is getting so caught up in designing "elegant" projects that they keep adding to and complicating their design to the point where it is indeed impressive—but impressive for its size rather than its usefulness. As you consider the various elements that will go into your design, include only those that directly address your current question. Make this project result in *useful* information within a relatively short time, and you'll probably find more support to set up future projects to answer some of the interesting questions you do not address this time. Don't try to squeeze too much into one evaluation!

Now we're ready to consider how to design your project.

#### **WHAT IS RESEARCH "DESIGN"?**

Research design, very simply, is the *recipe* you put together to answer your research questions. This step of the project is crucial. Just as a good recipe for chocolate cake can produce a pretty reasonable cake even if you don't have the finest of ingredients, a good research design can produce useful information even if you are stretched for resources and make a few mistakes along the way. On the other hand, the finest ingredients on earth will not produce much of a chocolate cake if you use a bacon omelet recipe, or have no recipe at all.

#### **DEFINING THE RESEARCH QUESTION**

The first step in the research design process is properly defining your research question. As you recall from our discussion in Chapter 1, there are many purposes for which people might do program evaluation research. Summative research asks questions about how well things are working and formative evaluation is to find out how things can be improved. The different levels of research we discussed (volume, reach, effect, and value) each lend themselves to different types of questions as you saw in the examples. However, regardless of what type or level of research you do,

you must still take the questions you ask and turn them into terms that can then be *measured*. Look at the following examples to see the difference between researchable questions and questions that are not (yet) research questions.

<b>DIFFERENCE BETWEEN GENERAL QUESTIONS AND RESEARCH QUESTIONS</b>	
<i>GENERAL QUESTION</i>	<i>RESEARCH QUESTION</i>
Do booster sessions in middle school improve D.A.R.E. outcomes?	Do booster sessions during the 1 <sup>st</sup> , 3 <sup>rd</sup> , and 6 <sup>th</sup> months of Grade 6 reduce the frequency of experimentation with controlled substances during that academic year?
Can comparable results be obtained with the low-cost drug, Makimwell, and the more expensive drug, Chargimwell?	What difference in number of cravings and relapse episodes are reported between individuals using Makimwell and those using Chargimwell?
Does the introduction of an anti-drug campaign in the community reduce drug abuse?	Does the XYZ anti-drug campaign reduce the number of first-time users within the 12 months following its introduction?
Does Treatment A work?	What changes are reported in clients' use of alcohol (other drugs) and happiness three months after they complete Treatment A?

One obvious difference between general questions and research questions is that the latter are longer. The *reason* for their greater length is that the researchable ones are also *more specific*. Questions that are too general such as "Does the treatment work?" do very little to help you design your project. In fact, such questions primarily generate more questions (which treatment, what do you mean by "work," for which clients, etc.). Good research questions state as exactly as possible what you plan to do (what treatment, program, change), and what will be measured.

Use Worksheet #7 on page 48 to refine the evaluation question you outlined in Chapter 2 into a researchable question. If you discover in the process that you are missing some important pieces, that's fine. In the chapters ahead we will give you more ideas about how and what to measure, how to follow up and for how long, etc. For now, make your question as specific and researchable as you can. From there, we'll design a project to answer your question.

**WORKSHEET #7. DEFINE YOUR RESEARCH QUESTION**

*Step 1: Copy your question from Worksheet #4.*

*Step 2: State what information you are looking for.*

Examine your question to make sure that a stranger could tell what it is you are looking for. For instance, a question that asks “How are we doing” does not offer much in the way of information about what you are looking for. It could as easily mean “how is the financial health of the organization?” as it could mean “do clients who complete the program seek further professional help during the next 12 months?” Phrase the question so that it clearly defines what you are going to measure. You might not yet know what measurement tools (surveys, questionnaires, etc.) you will use but you should know what information you are seeking. Refine the question you recorded in Step 1, by making it as specific as possible:

*Step 3: If you are evaluating a change to your program, define it in the question.*

The research question you define will drive the design of your project. If the question is vague about what you are evaluating, it will be difficult to put together a design that results in useful information. If you plan to make changes, examine your question again and make sure it specifies what you are going to do. If not, revise it here.

This page intentionally left blank.

By now your research question should be clear enough that you have some idea about how you might go about testing it. However, before plunging ahead take the time to make sure you have a good grasp of the basics of research design. You will probably find that much of what we have to say on this topic is either familiar or intuitive to you. Nonetheless, take the time to think through the issues. The more time you invest in careful design of your project, the less work it will involve down the road and the more valuable will be its results.

## **A FEW RESEARCH TERMS**

If you were to spend an afternoon telling a group of mechanical engineers who had no experience with the substance abuse field—either professionally or personally—about what you do, you might be astounded at the number of field-specific phrases you would need to stop and explain. People who work at length in a particular area tend to develop jargon that succinctly expresses their thoughts to others working in the same area. Although this sometimes makes it difficult for outsiders to understand the conversation, the jargon is usually quite useful in that it clearly defines the ideas being expressed. This is no less true for the area of research design than it is for other areas. Here are just a few basic research design terms that may be helpful to consider as you design your project.

### **Dependent Variables**

The most fundamental aspect of science is careful observation. In any evaluation project, you are observing what happens on certain specific dimensions. These might be behaviors (such as drug use, or coming back for help), attitudes, beliefs, knowledge, or other kinds of outcomes. These are things that change or vary over time and are therefore called *variables*. The variables you choose to watch—the ones in which you're interested in seeing a change, are known in research as the *dependent variables*. These dependent variables are the thing that (you hope) will change as a result of (*depending on*) what happens in your program. Here are some specific examples of dependent variables that are often used in studies of substance abuse treatment and prevention:

- ☞ Number of people who used (versus abstained from) drug X at least once during a certain period of time

- ☞ Percentage of days on which people were drug-free after the program
- ☞ Number, length, and severity of relapses
- ☞ Number of intravenous drug users who became HIV-positive
- ☞ Extent of positive versus negative attitudes toward drug X

In essence, dependent variables are the "things" you observe to tell you how your program is working. What are the "things" you care about most, the things that you hope your program will change? These are your dependent variables and should be reflected in the researchable question that you started to outline earlier in this chapter.

## **Data**

To evaluate whether your dependent variable changes, you must gather measurable information that tells you if it has changed and, if so, how much and in what direction. Complicated as this sounds, all "data" really means is the specific measurements you are collecting. For example, suppose you wanted to know whether a brief intervention, delivered to heavy drinkers being seen for routine medical care, would decrease heavy drinking during the next 6 months. The dependent variable you are watching is the frequency of heavy drinking but the actual data you gather might be questionnaire or interview scores that tell you about the person's drinking. ("Data," by the way, is a plural noun. One piece of data is called a datum.)

Data, in other words, are the currency of research. Without them, you have great ideas but no information (much like having a picture of the car you want but no money with which to buy it).

Before we leave the topic of data, we must emphasize the importance of being consistent in how you gather them. Even if you use the most objective and simplest of measurements it is important, as much as possible, to gather all the information in the same way. For instance, questionnaires administered to clients should be given in the same way and at the same point in treatment for all participants.

## **Independent Variables**

Often research involves making a change in one thing and seeing how it affects something else. Drug companies, for example, do studies to see how a new medication affects the symptoms and causes of the illness it is intended to treat. Not all research is like this. Astronomy, for example,

mostly involves observing how the universe works, without trying to change it. When you are studying the effects of changing something, you are studying the effects of an *independent* variable. The independent variable in a study is the "thing" that you control. This kind of research involves making a change in one thing (independent variable) to see how it affects something else (dependent variable). If you are evaluating a new anti-drug poster campaign for high-schools, the posters would be the independent variable because you, the researcher, control which posters go where.

Do you expect to have an independent variable in your evaluation project? Are you planning to control or change something to study its effect? Or do you plan to start just by studying dependent variables, without trying to change anything in the program? You might evaluate the effect of adding a new component (X) to the program, in which case X is your independent variable.

### **Predictor Variables**

You may wish to include variables of a third kind in your study. These are *predictor* variables. You don't necessarily expect them to change as you do dependent variables, and you are not changing them to see their effects as you would with independent variables. You are just interested in whether they *make a difference* in what happens. For example, you may wonder whether men and women experience different outcomes in your program. In this case, gender is a predictor variable. You are not controlling or changing gender, obviously, so it's not an independent variable in the experimental sense. Nevertheless you're interested in whether it makes a difference.

### **Control Group**

Suppose you offer a program for people who are convicted of drunk driving and you find that one year later 10% of those who came through your program have repeated their offense. Is that a good outcome or a poor one? Did your program have a beneficial impact or not? If you want to demonstrate conclusively that beneficial changes are really the *result* of your program, you will benefit from having a good control group. Whenever you evaluate a new or modified program component (your independent variable), you need a way of telling whether the effects you see are due to the your independent variable or due to some other factor such as the mere passage of time. You need to *control* other possible outside (extraneous) influences. The usual way to do this is by having two groups: one that receives the new and improved process you are testing (the experimental group) and one that does not (control group)—perhaps receiving the old approach without any of the changes. In every other way, except for the independent variable, the two groups should be alike. This kind of design tells you whether your program really made the difference. It is usually the best kind of evidence to have when you want to defend the effectiveness of your program.

### **Selecting and Assigning Participants**

If you are gathering data about *people*, those people are your research *participants*. Projects that ask volume and reach questions typically do not have specific participants because

they do not require much information about individuals. If your project plan includes getting data from individual participants, there are three issues you must consider.

- 1. Confidentiality and agreement to participate.** If you are asking people to participate in extra procedures that are not part of your regular program practices, you need to get each person's specific consent to participate in the project. Any institution that normally does research (like a university or hospital) has a formal procedure for gathering such consent. The basic procedure is to provide the person with a full and fair written description of what is involved in participating and have the person sign a statement agreeing to participate. This must include a clear description of confidentiality—who will have access to the information that is given. It should also include a clear statement that the person can refuse to participate, or decide later not to participate, without any penalty and still be entitled to the normal services of the program.
- 2. Sampling.** In a perfect world (and one with unlimited resources) you might want to include in your study every client who comes into contact with your program. However, we live in the real world and so it is seldom possible to include that many participants. If you must select a smaller group than *all* clients, it is important to select participants who are *representative* of the larger group. For example, if your practice is 85% female, try to ensure that your participant sample is not predominantly male. Otherwise, it will be difficult to conclude whether your results are applicable to your whole practice or just to the men. The simplest way to ensure that your sample is representative of the larger group is to pick participants *randomly*. Thus, instead of including in your study clients whom therapists think would be cooperative, pick your participants at random. If you are evaluating the effect of a school-based prevention program, gather your outcome data from a randomly selected sample of students, not only from those who are easiest to find or who were most actively involved in the program. Another possible way to get a reasonably representative but limited sample is to enter *everyone* into the study who enters the program during a certain period of time (one month, for example).
- 3. Assignment.** If you are studying the effects of an independent variable and will have at least two groups in your study (for example, experimental and control

groups), it is important to make sure that your groups are as similar as possible to begin with. If your experimental group consists of those who volunteered for a special program, and your control group consists of those who did not, you won't be able to conclude much at the end of your evaluation even if you observe a big difference in outcomes. That is why studies often assign people *randomly* to different groups—for example, by the toss of a coin. Another way to accomplish the same thing is to take people in the order that they come to your program and assign the first to group A, the next to B, the next to A, the next to B, and so on until you have enough people in both groups. This way you are not giving an unfair advantage to either group at the outset.

## **THE KISS OF USABLE RESEARCH**

There is one final reminder to consider before digging into designing your study. That is the KISS principle. Very simply: *Keep It Simple, Sweetie*. Simple project plans that clearly address the question you are asking and provide easily understood data are most effective.

Simple research designs and uncluttered data files are wise because they:

- ☞ Ensure the evaluator can understand his or her outcomes and consequently explain them clearly.
- ☞ Enable the audience to easily understand the findings and thus buy into the conclusions and support recommended program changes.
- ☞ Tend to require fewer financial and personnel resources and thereby are more affordable.
- ☞ Are less confusing to manage and therefore more likely to be followed through to a satisfactory completion.

As you design your project, keep asking yourself whether this or that addition really addresses your primary research question. If it does not, question the value of including it. You can't answer all of the interesting questions in one study, and evaluation projects are usually more rewarding when they are simpler and provide understandable findings within a reasonable period of time.

## A Case Example

To illustrate the ideas presented thus far and show how a simple design can yield important results, we offer a case example. This study was done in a real-life treatment agency in Albuquerque, and its results are reported in a scientific journal. It may help you to compare the complexity of your own design to this one as you go.

**STUDY:** Impact of Motivational Interviewing on Participation and Outcome in Residential Alcoholism Treatment by Janice M. Brown and William R. Miller. Published in *Psychology of Addictive Behaviors*, 1993, Vol. 7, No. 4, pages 212–218.

This study began with a problem posed by the clinical staff of a residential treatment facility:

"As health care funding tightens, we can only have people in the program for a short time, perhaps just two or three weeks. If they're not really motivated, it's hard to make enough progress with them in this short time. What could we do to motivate our clients, so that they are ready to work hard during the short time we have them with us?"

In response to this concern, the research team designed a two-session motivational intervention to be given immediately after people entered the residential treatment program. Here is how the study was designed:

The research question: Does the addition of a 2-session motivational interview upon admission to our residential treatment program for alcohol dependence improve treatment outcome?

Independent variable: *Experimental Group:* Given the new motivational interview and then the regular program  
*Control Group:* Given the regular program as usual, without the new motivational interview

Dependent variables: 1. Treatment participation as rated by therapists at discharge  
2. Amount of drinking at 3 month follow-up

Data (measures): 1. Therapist questionnaire asking about client's motivation and participation in the treatment program.  
2. Structured client interviews measuring alcohol consumption at 3-month follow-up

Participants: 28 consecutive admissions to the inpatient substance abuse facility between April and June 1991

Assignment:	Following admission and standard intake/assessment procedures (including alcohol consumption measures), participants were alternately assigned to either an experimental (motivational interview) or control (no interview) group. Participants in the experimental group received two interview sessions (with a structured interview protocol) prior to entering treatment. Participants in the control group went directly to treatment. Therapists were not informed of which clients had received the motivational interview.
Data collection:	At discharge, therapists completed questionnaires about the clients' participation in treatment.  Three months after discharge, all participants were contacted for a follow-up interview to assess alcohol consumption.
Data analyses:	Planned data analysis included comparison of program participation scores and changes in alcohol consumption between those receiving versus not receiving the motivational interview.
Findings:	Therapists perceived that clients from the experimental group were significantly more motivated and participated more actively during treatment.  Twice as many clients in the experimental group remained completely abstinent during the first three months after discharge  Drinking during the three months after discharge was more than three times higher among control group participants, who received the same inpatient treatment program but were not given the initial motivational interview.
Conclusion:	The addition of a relatively simple, low-cost motivational interview significantly increased client motivation. Therapists found the clients more rewarding to work with and treatment outcomes were dramatically improved. The two-session interview was judged to be a cost-effective addition to this 21-day residential program.

## ***BEGIN PLANNING YOUR PROJECT***

Worksheet #8 on page 60 will start you off in planning your evaluation project. At this point you have enough information to begin outlining your detailed research plan. You have already specified your research question and will now identify the dependent variables (and independent variables if you have them) as well as the general nature of the data you will gather. You can also get started on figuring out your *methods*, how you will gather the necessary information, but will undoubtedly need to adjust and adapt your methods as you pinpoint the specific measurement tools (Chapter 4) you will use as well as other considerations raised by your team members.

As you complete each of the remaining chapters in this manual, we will guide you in further refining your project design. By the end of the manual, you will have a complete research design and be ready to work out the administrative details with your team.

This page intentionally left blank.

WORKSHEET #8 . DESIGN YOUR STUDY .

FILL IN EACH SECTION WITH AS MUCH DETAIL AS YOU CAN.

Copy your *research question* from Worksheet #7. It is always best to have your objective clearly in front of you when designing the project meant to address it.

*Dependent variable(s)*. What do you expect to see changes in?

*Independent variable*. Is there anything you are going to control and change to study its effects?

*Predictor variable(s)*. Is there anything else you want to observe, to see whether it makes a difference in your dependent variable(s)?

*Data*. Although you have not yet selected your specific measurement instruments, you may have a general idea of the form your data will take, such as questionnaire scores or frequency information. If so, describe it now. What will you measure?

This page intentionally left blank.

## **SUMMARY**

Designing an evaluation project is much like putting together a recipe for a cake. It takes careful thought about what you want to accomplish (bake) and diligent attention to ensure that the methods (ingredients) you use are indeed appropriate for accomplishing your objective. Just as eggs, cheese, and bacon will not produce much of a chocolate cake, a hastily put together project plan may well produce a product, but not likely the one you want.

The components of research design are few but each is essential to consider. You must think about your *dependent variable(s)* (what you will observe to see whether it changes), any *independent variable* (something you will change and control, to see if it has an effect) or *predictor variable(s)* (something you don't directly control, but that may make a difference), the form of *data* you will gather, and the methods you will use to carry out your plan. Of course, having a plan for how you will analyze and use your data is also an essential part of designing a research project that will pay off in useful information. It is unwise to wait until your data have been collected before you consider what to do with them. We will consider issues of data analysis in more detail in Chapter 8. For now we turn our attention to the process of selecting the right measurement tools.

This page intentionally left blank.

## **4—MEASUREMENT TOOLS**

When you set out to evaluate a program, or anything else for that matter, you need measurement tools. You need to determine exactly what it is you are measuring and what units of measurement you are using. For instance, if you were teaching fifth grade physical education and wanted to know how effective the program was at increasing students' strength from the beginning to end of the first semester, you would need to measure *strength* and, consequently, need a tool for doing so. In this case, the "tool" would probably be a series of strength exercises the children could do to give you units of measurement such as repetitions (e.g., number of stomach crunches) and resistance force (e.g., number of pounds lifted).

In evaluating substance abuse programs, the tools you might use include feedback from community sources, program usage/reach numbers (e.g., traffic), clinical observation, structured interview data, client questionnaires, and biochemical markers of substance use (such as urine tests). Depending on the research question you ask, one, some, or all of these measurements may be of use. Each is valuable in its own way, and each presents potential pitfalls.

### **COMMUNITY FEEDBACK**

Research questions that ask how your program is perceived by the community, or that require input from members of the community regarding your clients, present you with a challenging task. You must figure out how to get *systematic* information from people who may randomly come and go throughout your day and have little reason to go out of their way for you.

If you are gathering information that reflects the community's *opinions* or *perceptions* about your facility, your best strategy for getting clear feedback is to develop a survey that asks specifically about your concerns. In so doing, it is important to keep in mind a few tricks for making sure the answers give you information that can be summarized, interpreted, and used. The following pointers apply whether your questionnaire is intended to be completed by the respondent, a face-to-face interviewer, or a telephone interviewer.

1. Open-ended opinion questions are terrific for putting your finger on the pulse of the community. If you survey 100 people and 70 of them spontaneously come up with the same or similar comments, you can be pretty sure you've hit a nerve. However, it is seldom the case that you find this strong a theme running through everyone's responses. More often, you end up with 100 disconnected thoughts that serve primarily to confuse those trying to make sense of them. That's why it is unwise to use *only* open-ended questions in an evaluation.

**Open-ended Opinion Questions**

- ☞ How do you feel about your experience with the program?
- ☞ What can we do to better meet the community's need?
- ☞ In what way did the program fail to meet your expectations?
- ☞ What were you expecting to gain from your involvement with us?

It is best to save your open-ended opinion questions for only one or two slots on your questionnaire and use them as you would a weather-vane; not to give you precise information but to help you detect an emergent trend *if one is indeed emerging*. Beyond that, stick to questions that generate answers that can be counted, summarized, and interpreted.

2. Open-ended questions that ask for specific numeric information work well if the answers are numbers that can meaningfully be averaged or grouped. For instance, if you ask each respondent for

**Open-ended Questions that Ask for Numeric Information**

- ☞ How old are you?
- ☞ How many times have you met with one of our counselors?
- ☞ How often have you seen a field rep from our agency on campus?
- ☞ How often do you eat in a restaurant that permits smoking?

his/her age, all the answers can be meaningfully averaged into a single number that describes the average age of your sample. The same is true of asking, for example, about how many days an abstinent client has thought of using his/her preferred substance in the past month. A different approach is to group

such numbers into meaningful ranges that would tell you what proportion of people fell into younger, middle-aged, or older categories, or reported never, infrequently, or frequently craving.

3. Multiple choice questions are ideal for surveys. Whether you ask yes/no questions that have only two possibilities or questions that list 13 program attributes that you want to know whether people find desirable, these questions give you clear answers that you can count, summarize, and interpret.

Keep in mind that you can also use multiple choice questions to get numeric information by pre-grouping the responses for your participants. The last question in the Multiple Choice Questions example box illustrates this.

**Multiple Choice Questions**

☞ Have you ever used our facility? Yes No

☞ Which of the following skills have you used during the past 30 days to help you resist drugs? (Check all that apply.)

- a) avoiding people who use drugs
- b) keeping busy
- c) going to Twelve Step meetings
- d) exercising
- e).....

What is the combined income for your household?

- a) under \$20,000
- b) \$20,001-\$40,000
- c) \$40,001-\$60,000
- e) \$60,001-\$80,000
- f) over \$80,000

## **PROGRAM USAGE/REACH NUMBERS**

A common mistake new researchers make is assuming that just because they are gathering numeric data like head-count or money, the information is by nature "pure" or "true." This is not necessarily the case. If you do not have a good plan for gathering your numbers, they may give you even less reliable information than if you just asked people for estimates. For example, let's say you want to know the volume of calls coming into your emergency line, so you count the calls received during one week. Based on the large number of calls received, your agency adds another person to call-duty—a person who ends up with little to do because, as it happened, you took your sample during the third week of December when the flow of calls is extraordinarily high. Had you *sampled* the call volume 2 or 3 days each week of the month over several months you might have realized that the

week you sampled was highly unusual in its volume, or that your third week of the month is always high and the rest of the month considerably lower. In this case, your staff could probably have given you a more accurate estimate of the actual flow.

When gathering numbers about program usage or cost, it is best to have a systematic plan. Think about whether the numbers are likely to be randomly distributed or clustered around one value range (such as the call frequency in the example above) and plan to gather them so that they are representative. Usually, you can accomplish this by gathering your information across an extended period of time or across a variety of settings.

## **CLINICAL OBSERVATION**

Another possible source of information about your clients is the people who work with them—therapists or prevention teachers, for example. Informal clinical observation can capture nuances of human behavior and emotion that are difficult to capture by questionnaire. Unfortunately, when you are trying to get a meaningful look at the big picture of a program, clinical observation is painfully insufficient. Aside from the obvious obstacles that get in your way such as deciphering practitioners' handwriting or wading through lengthy files, there are more serious obstacles such as the inconsistency among staff in how carefully they interpret diagnostic criteria, document change, describe severity, and so on. Experience also shows that those who deliver services are usually not the best judges of program effectiveness. If you are after information that can be pooled to reflect your overall program, clinical observations may not be the most reliable road to travel.

## **BIOMEDICAL MARKERS OF SUBSTANCE USE**

Depending upon what substances you are interested in monitoring and for what purposes, there are various biological markers that you can measure, such as breath, blood, or urine tests. One real advantage of such measures is their objectivity—they are hard to fake. If someone says they are not using cocaine but it shows up in their urine, you can believe the biomedical test.

Yet such tests are not without their shortcomings. It turns out that in many cases people will honestly report having used a drug that is not detected by a biomedical test. This

can happen because certain drugs do not stay in the body for very long. In other cases, a person may truthfully say that they have not used marijuana within the past two weeks, although residue from prior use still turns up in a urine sample. The collection of urine and breath samples is also rather intrusive and testing can be expensive. A common procedure these days, if you can manage the cost and practical problems of testing, is to use *both* client self-report *and* biomedical markers.

**Breath Tests.** Perhaps least intrusive of the biomedical tests are breath tests. These are particularly used to check for alcohol in the body and are as reliable as urine and blood testing as long as proper procedures are followed. Breath tests can be given routinely at the beginning of every session, or can be used only when there is suspicion of drinking. One of the more economical equipment models is the Alco-Sensor, a hand-held unit that costs about \$600 (plus minor cost for mouthpieces) and can be used about 500 times before it needs to be rebuilt at a cost of about \$150<sup>4</sup>. Breath tests for thiocyanate have also been used to confirm self-reports of abstinence from tobacco smoking.

**Urine Tests.** Urine drug screening (UDS) is increasingly common as employees in critical occupations are required to submit to random testing. UDS is also commonly used in drug treatment centers (such as methadone maintenance programs) to confirm abstinence from illicit drugs. A full screen (usually done by a procedure called Enzyme Multiplied Immunoassay Technique, or EMIT) can be done for the presence of metabolites of most drugs of abuse, though there is presently no good test for the use of inhalants, for example. When certain drugs (particularly amphetamines) are detected in a UDS, it is necessary to conduct a more thorough procedure to confirm actual use because other substances can cause a positive screen. The cost of these second-pass confirmations is usually built into a lab's charge for UDS. Because there are ways to fake a urine test (for example, to insert water or someone else's urine into the container), various levels of monitoring of sample collection are used, depending upon how the samples will be used. Less intrusive checks are to ensure that the sample is at body temperature and has the right pH value for urine. Other programs have a staff member directly observe collection of the sample. In order to reliably ensure drug abstinence, it is necessary to collect UDS samples 2-3 times per week, because many drugs are eliminated rapidly from the body. UDS can cost between \$12 and \$50 per test, depending upon the lab and procedures used.

---

<sup>4</sup> Ordering information for the Alco-Sensor is given in Appendix A.

**Blood tests.** Blood samples can also be tested for the presence of alcohol and other drugs. In addition, blood samples can be used to test for other medical indications of drug (particularly alcohol) use. This is a possible advantage in blood testing, because medical indicators are longer-lasting than the presence of drugs in the blood system. Such measures can also be useful in monitoring the health of clients. Liver function tests such as SGOT and SGPT may reflect the impact of alcohol on the liver, and gamma glutamyl transpeptidase (GGT) is particularly sensitive to recent binges. There is increasing interest in carbohydrate deficient transferrin (CDT) as a sensitive indicator of recent alcohol use in males, although there may be problems with its specificity in women. Mean corpuscular volume (MCV), uric acid, bilirubin, and triglycerides and cholesterol are also commonly elevated in heavy drinkers. High density lipoproteins (HDL), the so-called "good cholesterol," tends to be elevated (which is desirable) by moderate drinking, but usually drops into a higher-risk range with heavy drinking. The cost of blood tests will depend upon the lab and number of tests used.

## **STRUCTURED INTERVIEWS & QUESTIONNAIRES**

Structured interviews and published questionnaires are perhaps the easiest and most economical way to get reliable information about your program. Using standardized instruments avoids most of the subjectivity involved in community feedback and clinical observation and improves the consistency of measurement from one client to the next.

There are many structured interviews and client questionnaires. Several excellent handbooks and summaries of these have been published and are listed at the end of this chapter. The instruments you choose will depend upon their purpose—whether you plan to use them for screening, diagnosis, understanding your client's problems and need for treatment, or follow-up. In addition, you should consider the extent to which the instrument has been carefully evaluated. We specifically recommend that you *do not* make up your own instruments for screening, diagnosis, assessment, or follow-up. Developing a reliable and valid instrument is a difficult process and there are now many good options available to you.

When you choose the instruments to use in your program evaluation, you should consider the following:

- Does it really measure what it says it measures, and what you want to know? Just because an instrument is called a dependence scale does not mean that it

accurately measures dependence. First figure out what you want to measure, and then search for research-based instruments that validly measure what you want to know.

- Does it work well for the population you are interested in? It is not safe to assume that an instrument developed for one group or culture will be accurate when applied to another group. This issue is considered in more detail in Chapter 5.
- Does it require specialized training to use? Some instruments require extensive and specialized training. Some may be sold only to properly qualified and trained professionals. Can it be used by those in your program who would administer it?
- Can it be easily scored and interpreted? Some instruments must be computer administered, or require complicated scoring procedures. Others come with easy-to-use scoring keys and can be scored accurately by clerical staff. Are norms provided in the test manual to help you interpret what the scores mean?
- Has this instrument been carefully tested? For well-developed instruments, the test manual will describe the research procedures used to validate the test. Is the test reliable—does it give the same results when administered twice, or by different interviewers? Does it measure what it is purported to measure, and not something else (like reading ability, general intelligence, or social desirability)?
- Is it affordable? Some tests are developed "in the public domain" (e.g., through federal research funds) and can be copied and used without charge. Copyrighted tests must be purchased and it is illegal to make your own copies instead of buying them from the publisher. Some tests are reusable, and for others you must buy new materials for each administration.

Later in this chapter we will make a few specific recommendations for reliable instruments that can be used easily and inexpensively in most substance abuse program settings. Copies of some of these instruments that are in the public domain are included in Appendix A. Other resources are listed at the end of this chapter. Before we get into the specifics of various measurement tools, however, there are three concepts related to test design that are important for you to understand: reliability, validity, and interpretability.

## **RELIABILITY**

*Reliability* refers to whether the instrument gives the same results each time it is used with the same person (who has not changed since its last use). This is easier to understand if you think in terms of physical things. Consider a thermometer that indicates you have a fever of 105° F and then, 1 minute later indicates your temperature is at 98.6° F. Assuming that you have not discovered the miracle fever cure, these readings tell you that the thermometer is not a reliable instrument. If you use an instrument to evaluate individuals and the instrument is reliable, it should provide basically the same information if it is given twice to the same person, or if two interviewers do the same interview with the same client. The latter is known as *inter-rater reliability*. An instrument with good inter-rater reliability will give the same results no matter who uses it (assuming they are properly trained). Inter-rater reliability is usually stronger for instruments that depend less on rater interpretation and more on objective information.

## **VALIDITY**

At the very least, instruments that are used in clinical work and in program evaluation should be *reliable*. It is possible, however, for an instrument to be reliable but not valid. Suppose you were to measure how high a person can jump and use this as an indicator of intelligence. The measurement would be reliable, in that you would get about the same result each time you asked the person to jump as high as possible. Yet this is not a *valid* indication of intelligence.

To say that a measurement instrument is *valid* means that it accurately measures what it is supposed to measure. Here are some different technical meanings of the term "validity" when applied to a measurement tool.

### **Face Validity**

*Does the test look like it measures what it claims to measure?* This is the loosest form of validity. The term "face" validity comes from the fact that one is looking at the test superficially—only at its face—and not beneath the surface. On the face of it, you might expect a screening test for alcoholism to ask about things like memory blackouts and hangovers, and not about which President the person

believes was the greatest. It is possible, of course, to have a subtle or "sneaky" test that asks indirect questions. Such a test would not appear to be measuring what it is supposed to measure, and so would have low face validity. Yet, in other senses discussed below, it might be valid.

### **Convergent Validity**

*Does the test give the same results as other tests of the same thing?* Going deeper than face validity, one can ask whether the results of this test converge with the results of other tests that are already known to be accurate. Are the results similar? Would one come to the same conclusion or diagnosis? In essence, this form of test evaluation involves comparing the test in question with a known standard. Turning again to the physical world for easy examples, consider a pregnancy test that fails to show positive when the woman's physical signs clearly indicate she is pregnant. This pregnancy test would not show convergent validity with the physical standard. When new pregnancy screening tests are being evaluated, they are compared for accuracy against other pregnancy tests of known accuracy.

### **Divergent Validity**

*Does the test measure only what it is supposed to measure?* Tests that are designed to measure one particular thing should not also be measuring something else. For example, a test that is supposed to measure self-esteem should not be strongly correlated with measures of intelligence. If it is, the test is mixing self-esteem with intelligence and thus makes it difficult to know just what the test score is telling you.

### **Predictive Validity**

*Does the test make accurate predictions?* For example, if a test is supposed to predict the likelihood that a drunk driver will repeat the offense, how accurate are its predictions? Not all tests, of course, are meant to predict the future.

## **INTERPRETABILITY**

All the scores in the universe are of no value if you cannot interpret their meaning. In working with measurement instruments, it is usually the case that you need some outside standard against which to interpret your results. For example, knowing that someone has scored 100 on an intelligence test tells you nothing unless you know that the

test is designed so that 100 is the population average. Accounting data are more useful if you can compare your net revenues for this month with the average monthly net revenue this year or with this same month last year. Being able to interpret your data relative to a known standard moves you from the position of having information to having *meaningful* information.

Many psychological instruments include *normative standards* to help users interpret the test scores. Normative standards are basically tables of scores that allow you to compare your participants' scores with those of a larger reference group. By comparing participants' scores with the norms, you can more accurately tell what the obtained scores mean. In selecting normed instruments for use in your study, keep in mind that norms are specific to the population on which they are based. If the test was normed on adults, for instance, comparing adolescents' scores to the norms will not necessarily be meaningful.

Another way that test-makers help to make their tests interpretable is by including subscales. Along with the overall scores, subscale scores can provide more specific information about the behaviors, symptoms, or problems associated with the total score. Some instruments can be scored to give you a *profile* of the individual across various subscales, which can be helpful in treatment planning or just in getting a more complete picture of the person's status.

## **RECOMMENDED INSTRUMENTS**

What instruments are best for your own program evaluation? The tools you select should be reliable. They should be valid for the purposes to which you will put them. They should give you interpretable information that you can use. There are literally hundreds of instruments in the alcohol/drug field from which to choose. Here are some recommendations of carefully developed instruments that have been shown to be reliable, valid for certain purposes, and for which interpretation is guided by clear norms or profiles. Wherever possible, we have described instruments that were developed in the public domain, which means that they can be reproduced for your use free of charge. Other instruments listed here are copyrighted, which means that individual copies must be purchased and it is illegal to make photocopies. Appendix A contains copies of most of the public domain instruments we discuss and ordering information for the balance. It also provides sources from which you may purchase the copyrighted materials. These are only a few of the instruments available. For more options, see the resources listed at the end of this chapter.

### **Screening**

Screening instruments are used to determine whether or not a substance use disorder is present. Given that it is usually best to err on the side of safety, these tests are designed to be more rather than less sensitive. Thus, they are more likely to detect a usage disorder in someone who actually does not have a problem than overlook one in an individual who does. As a result, a positive outcome should always be followed up with more careful diagnosis and assessment.

***Alcohol Use Disorders Identification Test (AUDIT)***<sup>5</sup>. This 10-item screening instrument is publicly available and takes only a few minutes for the client to complete. It is also simple and quick to score by hand and has cross-cultural validity. The range of possible scores is zero to 40 with a score of 8 or higher indicating the likelihood of hazardous drinking and the need for further assessment. A copy of the instrument and scoring instructions are included in Appendix A.

---

<sup>5</sup> Saunders, J.B., Aasland, O.G., Babor, T.F., DeLa Fuente, J.R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative project on early detection of persons with harmful alcohol consumption-II. *Addictions*, 88, 791-804.

**Drug Abuse Screening Test (DAST)<sup>6</sup>.** Self-administered in about 5 minutes, the DAST contains 20 yes/no questions about drug use over the prior 12 months. A score of 6 or more indicates the presence of a moderate (or more severe) degree of problems related to drug abuse. The DAST is copyrighted. Ordering information appears in Appendix A.

## Diagnosis

When your goal is to establish a diagnosis, it is best to stay very close to the criteria found in the diagnostic system that you use. The following structured interviews provide diagnoses that correspond to the Diagnostic and Statistical Manual (DSM) that is commonly used to classify psychological disorders including substance abuse.

**Structured Clinical Interview for DSM, Patient Edition (SCID-P).** Developed by Spitzer, Williams, Gibbon, and First<sup>7</sup>, this is probably one of the most popular DSM interviews. The SCID-P provides diagnoses for a range of clinical disorders yet is structured to allow you to administer only the alcohol and drug questions if that is what you need. It is copyrighted, however, so the entire instrument must be purchased regardless of the questions you intend to use.

The Alcohol section of the SCID-P takes approximately 15-20 minutes to administer and generates a diagnosis of Alcohol Abuse or Alcohol Dependence with dependence being classified as mild, moderate, or severe, or as being in partial or full remission.

**Diagnostic Interview Schedule (DIS).** This interview was developed by Robbins, Cottler, & Keating<sup>8</sup> at the request of the National Institute of Mental Health. The request came in response to the need for a diagnostic interview that could be easily administered by lay personnel. Accordingly, the DIS is an easily scored interview that can be administered by non-clinical personnel given adequate training. A computerized version of the DIS (the C-DIS) is also available.

As is the case for the SCID-P, the DIS addresses a range of diagnostic categories but is structured to allow you to administer the alcohol and drug section independently. The

---

<sup>6</sup> Skinner, H.A. (1982). The drug abuse screening test. *Addictive Behaviors*, 7, 363-371.

<sup>7</sup> Spitzer, R.L., Williams, J.B.W., Gibbon, M., & First, M.B. (1990). *Structured Clinical Interview for DSM-III-R-Patient Edition (SCID-P, Version 1.0)*. Washington, DC: American Psychiatric Press.

<sup>8</sup> Robbins, L., Cottler, L., & Keating, S. (1989). *The NIMH diagnostic interview schedule. Version III, revised (DIS-III-R)*. Rockville, MD: National Institute on Mental Health.

alcohol section consists of 30 questions, some with sub-questions, and are read by the interviewer exactly as written. Clients typically have little trouble understanding the questions. The alcohol section of the DIS takes approximately 15-20 minutes to administer. The paper and pencil version of DIS is not copyrighted and the computerized version is. Ordering information for both versions is provided in Appendix A.

## Assessment

So far we have talked about instruments to screen for the presence of a problem and diagnose the nature of that problem. In a sense, both of those sets of instruments tell you more about the problem than they do about the individual client. Assessment instruments are designed to tell you about the individual him- or herself. These instruments help you develop a picture of the client's unique situation as it relates to his/her substance use.

In assessing substance use, it is necessary to look at more than just whether or not the individual uses. Obviously, degrees and patterns of usage will make a difference in how you treat the individual as well as in what the prognosis and some of the road blocks are likely to be. Thus, assessment should occur in a number of areas. You will want to know how much and how often the person is drinking/using (patterns of use). It is also important to identify what life problems (e.g., financial strain, loss of employment, legal hassles, marital stresses) are caused by or related to his/her usage patterns. Finally, you will need to know just how severe the individual's dependence really is.

## Patterns of Use.

The simplest means of determining usage patterns is to ask the individual how much (quantity) and how frequently s/he uses. These Quantity/Frequency (QF) Questionnaires typically take the form of two simple questions as shown here and require only a moment to administer. Depending on the drug about which you are asking, it may be easier or more difficult to specify a quantity of use. For example, a standard drink unit is typically defined as 10

### **Sample QF Questions**

- ☞ During the past \_\_\_\_ days, on how many days did you drink any alcohol?
- ☞ On days you did drink, how many drinks did you usually have?

ounces of beer, 4 ounces of wine, or 1 ounce of liquor or other distilled spirits such as gin and whisky.

It is helpful to define what constitutes a "unit of use" if you are using a QF Questionnaire to determine usage of other drugs. However, doing so is more complicated than it is with alcohol and cigarettes (which come in neat, government-prescribed strengths and packages). Street drugs vary in strength and composition and the user rarely knows for sure exactly what is in the substance. Nonetheless, an attempt should be made to arrive at a unit definition that is meaningful to both you and the client.

Appendix A includes three QF instruments used at CASAA that are not copyrighted and may be copied for use with your clients. The **QFV-30 Questionnaire** and **QFV-90 Questionnaire** are simple 3-question forms that query clients about their alcohol consumption during the prior 30- or 90-day period. If you choose to use this instrument as part of your follow-up evaluation, pay attention to the time span between administrations. If you compare its results at any point in time with its results less than 30 days after that point, the information you receive will overlap and your results will be difficult to interpret.

The third QF instrument in Appendix A is the **Quantity/Frequency Interview**. It asks for frequency and quantity of use during the past 30 days across a number of drugs. As you go down the list and record quantity, remember to clarify with the client what amount is meant by "uses."

Also useful for gathering usage pattern information, as well as other valuable information, is to ask clients to keep a **diary** of their intake. An advantage of this approach is that you do not need to rely on memory and the disadvantage, obviously, is that clients may have difficulty or be reluctant to keep a diary.

If you can get clients to self-monitor, however, you not only know how much and how frequently they are

<b>Sample Diary</b>			
Date	Time	Situation	Type and amount
_____			
_____			

using, but you can begin to identify situations that are particularly high-risk as well as over what situations mastery gradually emerges. A sample diary format is shown.

## **Substance Use Related Problems.**

Along with knowing how much and how often an individual drinks or takes other drugs, it is important to have a picture of how this usage affects his/her life. Excessive use of alcohol and other drugs can have negative effects across a spectrum of physical, social, and psychological domains. Even once usage has been curtailed or stopped altogether, these other problems may persist and make maintenance of the improvements all that much more difficult. Understanding how substance use affects the individual's life is a critical step in helping that person to make positive changes.

The next 4 self-administered questionnaires report on problems related to alcohol and general drug use. The fifth questionnaire under this heading is one you complete with the client. It gives you a sense of where the client believes his/her greatest problems lie. All 5 are provided in Appendix A.

***Drinker Inventory of Consequences—Lifetime (DrInC-2L)***<sup>9</sup>. This 50 item questionnaire asks whether 50 different problems have ever occurred during the client's lifetime. It has five subscales of 7 to 12 questions each. They are: physical, social responsibility, intrapersonal, impulse control, and interpersonal. The DrInC is easily scored by hand and comes with separate interpretation profile sheets for men and women.

***Drinker Inventory of Consequences—Recent (DrInC-2R)***. This is an identical copy of the DrInC-2L except that it queries the presence of problems within the past 3 months. It is also easily scored and comes with separate profiles for men and women to help you interpret the results.

***Inventory of Drug Use Consequences—Lifetime (InDUC-2L)***<sup>10</sup>. Similar to the DrInC-2L, the InDUC-2L asks about the occurrence of 50 problems over the life of the individual. It, too, provides 5 subscales (physical, interpersonal, intrapersonal, impulse control, and social responsibility) and is easily scored by hand.

---

<sup>9</sup> Miller, W.R., Tonigan, J.S., & Longabaugh, R. (1995). *The Drinker Inventory of Consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse. Test manual.* (Volume 4, Project MATCH Monograph Series). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

<sup>10</sup> Miller, W.R., Tonigan, J.S., & Longabaugh, R. (1995). *The Drinker Inventory of Consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse. Test manual.* (Volume 4, Project MATCH Monograph Series). Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

***Inventory of Drug Use Consequences—Recent (InDUC-2R)***. This is the InDUC version that asks about occurrence of 50 problems within the past three months. It provides the same subscales and scoring guide as the InDUC-2L.

The *Lifetime* (L) versions of these instruments are usually used just once, at first contact, to get a picture of lifetime severity. If only recent problems are of interest, the L versions are not used. The *Recent* (R) versions can be given both before and after program involvement and compared to show change. Instructions could be changed to ask about a period of time longer or shorter than 3 months. The DrInC-R norms, however, are based on the 3-month instructions.

***Goal Attainment Scale (GAS)***. In this simple questionnaire based on a method devised by Thomas Kiresak, the client defines one to three problem areas that s/he wants to address. For each problem, the client describes its current status and what deterioration and improvement would look like. Unlike the questionnaires described above, the GAS gives you a direct line to what the client sees as his/her greatest difficulties with respect to substance abuse. It is also particularly good for follow-up since it allows both client and therapist to easily identify achievement of positive and negative milestones. A further advantage is that it is individualized to each participant.

This simple instrument focuses on specific behaviors the client tells you would indicate progress. Because the test wording does not include any abstract concepts, it can probably be translated into other languages with little risk of miscommunication.

### **Dependence Measures.**

It is not always the case that someone who abuses a substance (and suffers negative consequences), is indeed addicted or dependent on the drug. Knowing whether the client is dependent on the substance in question does, of course, have a bearing on how you handle that individual.

The two alcohol dependence measures that we recommend have been well-tested and can be self-administered by clients in under 10 minutes.

***Alcohol Dependence Scale (ADS)***<sup>11</sup>. This 25-item questionnaire investigates dependence symptoms during the past 12 months. It is easily scored and comes with an excellent test manual that provides normative data against which to compare your

---

<sup>11</sup> Skinner, H.A. & Horn, J.L. (1984). *Alcohol Dependence Scale (ADS) User's Guide*. Addiction Research Foundation, Toronto.

clients' scores. The ADS is copyrighted. Ordering information appears in Appendix A.

**Severity of Alcohol Dependence Questionnaire (SADQ)<sup>12</sup>.** This 20-item questionnaire focuses on a 30-day period of heavy drinking and includes 5 subscales. They are: physical withdrawal, affective withdrawal, withdrawal relief signs, consumption quantity, and reinstatement rapidity. The test is easily hand-scored and includes guidelines for interpreting its results. It is publicly available and a copy may be found in Appendix A.

## **Comprehensive Instruments & Interviews**

The more comprehensive your picture of an individual's alcohol/drug "life," the more appropriate will be your treatment plans and the more information you will have with which to assess change. Fortunately, there are a number of instruments that provide comprehensive assessment across a variety of domains.

The two comprehensive instruments we recommend here are among the strongest available. One, the Alcohol Use Inventory, is a paper and pencil questionnaire that clients complete themselves and the other, the Addiction Severity Index, is a structured interview.

**Alcohol Use Inventory (AUI)<sup>13</sup>.** The AUI's 228 items yield a profile of the individual that provides scores along 24 dimensions including motivation for and styles of drinking, physical dependence, loss of behavioral control, relationship of drinking to marital problems, readiness for change, and others. An overall score for severity of alcohol involvement is also provided.

The AUI can be hand scored in a few minutes by a clerical worker or scored by computer. The manual provides clear guidelines for scoring and interpretation as well as normative data. The test is copyrighted and so cannot be reproduced in this manual. Ordering information is given in Appendix A.

**Drug Use Self Report (DUSR)<sup>14</sup>.** The DUSR measures drug use in 10 drug use categories and provides scales pertaining to drug

---

<sup>12</sup> Stockwell, T.R., Hodgson, R.J., Edwards, G., Taylor, C., & Rankin, H. (1979). The development of a questionnaire to measure severity of alcohol dependence. *British Journal of Addiction*, 74, 79-87.

<sup>13</sup> Horn, J.L., Wanberg, K. W., & Foster, F.M. (1987). *Guide to the Alcohol Use Inventory*, National Computer Systems, Minneapolis, MN.

<sup>14</sup> Wanburg, K.W. & Horn, J.L. (1989)/. *Guide to the use of the Drug Use Self-Report*. Arvada, CO: Center for Addiction Research & Evaluation.

use symptoms, benefits from drug use, and psychosocial functioning. Ordering information is provided in Appendix A.

**Addiction Severity Index (ASI)**<sup>15,16,17</sup>. This structured interview takes about 40 minutes to complete and is applicable to both primary alcohol and other drug areas. It provides scores for 8 subscales: life problems, medical, legal, employment/support, alcohol, other drug, family/social, and psychiatric functioning. It also asks the interviewer to rate the severity of each area along with his/her confidence in the client's truthfulness and ability to understand. The ASI comes with an administration manual that offers instructions for scoring and interpreting each area evaluated. The ASI can easily be used for follow-up as well. The printed form itself flags for you which questions to re-ask at follow-up. Ordering information is in Appendix A.

## **INSTRUMENTS FOR FOLLOW-UP EVALUATION**

Chapter 6 discusses the ins and outs of follow-up in detail. At this point, however, we want to say a few words about instruments that are appropriate for follow-up.

You can usually use the same instruments for follow-up as you do for your initial and post-treatment assessments providing the timeframe queried makes sense. For instance, the DrInC-2R would make sense both at intake, post-treatment, and follow-up because it asks about drinking within the recent past. However, the DrInC-2L makes sense only at the beginning of treatment and not the end because it inquires about lifetime drinking consequences. (This is obviously because treatment cannot change the client's earlier lifetime experiences.) Additionally, we might add, even the DrInC-2R makes sense at post-treatment and follow-up only if 3+ months have elapsed since its last administration.

Few instruments are developed exclusively for follow-up. However, you can conduct valuable follow-up evaluation by applying a few basic rules to selection of your instruments—

---

<sup>15</sup> McLellan, A.T., Luborsky, L., O'Brien, C.P., & Woody, G.E. (1980). An improved evaluation instrument for substance abuse patients: The Addiction Severity Index. *Journal of Nervous and Mental Disease*, 168, 26-33.

<sup>16</sup> McLellan, A.T., Luborsky, L., Cacciola, J., Griffith, J., Evans, F., Barr, H.L., & O'Brien, C.P. (1985). New data from the Addiction Severity Index: Reliability and validity in three centers. *Journal of Mental and Nervous Disease*, 173, 412-423.

<sup>17</sup> McLellan, A.T., Parikh, G., Bragg, A., Cacciola, J., Fureman, B., & Incmikofki, R. (1990). *Addiction Severity Index administration manual* (5th ed.). Philadelphia: Penn-VA Center for Studies of Addiction.

even if the instruments were not designed expressly for this purpose.

- Select an instrument that measures what you want to change with your program. In particular, ask yourself what long-term (post-program) changes do you want to see occur and choose your instrument accordingly.
- If you need to change the timeframe queried by the instrument, do so cautiously. Although you can often make this change without harming the integrity of the instrument, you will in almost all cases invalidate any norms tables that come with it.
- It is best if you ask about the same timeframe each time you assess clients. For instance, if your intake evaluation asks about substance use during the prior 30 days, your post-program and later follow-up evaluations should also ask about the prior 30 days.
- Use the same or parallel measures for pre- and post-program assessments. Remember that your objective in follow-up evaluation is to measure change over time. To do so, you need a common measurement tool at your different points in time so that you can compare results. Some instruments have parallel versions (that is, different-looking versions of the same instrument) that allow you to use the same instrument twice without the client being aware you are doing so. More often, however, you will need to simply use the same instrument at both (or more) times. In substance abuse evaluation, clients' awareness of this is typically not a problem.

## **SUMMARY**

Give as much careful consideration to selection of your measurement tools as you do to your project design. There are so many instruments available (some of them good, some of them not) for so many purposes that it is easy to get lost in the woods. Indeed, it is easy to create your own confusing tangle of woods by getting carried away with your choices and using too many instruments in the same study. Too much of even a good thing is not so good. Too many measures serve, in the end to confuse rather than clarify. Select carefully and only what you need.

The instruments we have recommended are reliable ones that have been evaluated for validity and are reasonably easy to interpret. There are certainly other good instruments available. Making sense of the test evaluation literature is

not a simple task and we recommend that you seek advice from someone trained in this area (usually a clinical psychologist) if you need assistance.

Before moving forward, take a moment to review Worksheet #8 and decide which instruments will best measure the dependent variables you are interested in. Record these on Worksheet #9 (page 85) and be prepared to adjust your plans in light of the discussions that follow in Chapters 5 and 6. Your project will become more refined as your knowledge grows.

This page intentionally left blank.

WORKSHEET #9. SELECT MEASUREMENT INSTRUMENTS.

Review your study design in Worksheet #8 and select the instruments that will best measure the variables in which you are interested. You will not necessarily need instruments from each category below.

***Screening Instruments***

***Diagnostic Instruments***

***Assessment Instruments***

***Comprehensive Instruments & Interviews***

This page intentionally left blank.

## **ADDITIONAL RESOURCES**

Here are additional resources where you will find good summaries or collections of measurement instruments that may be useful in your program evaluation.

Addiction Research Foundation. (1993). *Directory Of Client Outcome Measures For Addictions Treatment Programs*. Toronto: Addiction Research Foundation.

Donovan, Dennis M. & Marlatt, G. Alan (Eds.) (1988). *Assessment of Addictive Behaviors*. New York: Guilford Press.

Litten, R.Z. & Allen, J.P. (Eds.) (1992). *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*. Totowa, NJ: Humana Press.

Miller, William R., Westerberg, Verner S., & Waldron, Holly B. (1995). Evaluating alcohol problems in adults and adolescents. In Reid K. Hester and William R. Miller (Eds.) *Handbook of Alcoholism Treatment Approaches: Effective Alternatives*, pp. 61-88. Boston: Allyn and Bacon.

National Institute on Alcohol Abuse and Alcoholism (in press). *Handbook of Alcoholism Assessment Instruments*. Rockville, MD: NIAAA. For a free copy call The National Clearinghouse for Alcohol and Drug Information. (1-800-729-6686).

This page intentionally left blank.

## **5—ASSESSING SPECIAL POPULATIONS**

When living in a culture as diverse as many parts of America, it is responsible to ask what special sensitivities or issues are needed when working with specific subgroups. How are such considerations of special populations relevant for the purposes of program evaluation?

First, we should consider what is meant by "special." One possible meaning is just "specific"—a group with certain defining characteristics. Another meaning distinguishes special populations from a "majority" group, the latter often thought of in America as white, English-speaking adults. In the American Southwest, this group is often referred to as "Anglos," and though the term is less than accurate, we will use it in this chapter for convenience of description. Although Anglos may not make up the majority of your own client population, it has been the group used to develop most commonly used measurement instruments. Thus, in selecting instruments for other "special" populations, there are a number of important issues you need to consider.

### **AGE**

Most of the substance abuse instruments currently in use were developed with an adult population and have not yet been properly evaluated with children and adolescents. It may be tempting to say that "abuse is abuse" and apply the instruments across age groups, but there are differences in how younger people use and respond to alcohol and other substances. For example, most adolescents who drink do so sporadically and their use of alcohol does not resemble the steady maintenance drinking of adult alcoholics. They typically show fewer symptoms of alcohol dependence, which is a key diagnostic concept in working with adults. Thus, using instruments that have not been evaluated specifically with the age group you are working with can provide you with misleading information.

The following table lists some instruments that have been validated with young populations. For more specific information regarding the assessment of adolescent substance abuse, consult the chapter by Miller, Westerberg, and Waldron (1995) listed at the end of Chapter 4.

<b>Assessment Instruments For Use With Adolescents</b>		
<i>INSTRUMENT/ CITATION</i>	<i>DESCRIPTION</i>	<i>AVAILABILITY</i>
<p>Adolescent Drinking Index (ADI)</p> <p>Harrell, A.V. &amp; Wirtz, P.W. (1989). Adolescent Drinking Index test and manual. Odessa, FL: Psychological Assessment Resources.</p>	<p>24-item questionnaire screens for alcohol abuse by assessing loss of control, and social, psychological, and physical indicators; reports subscores for self-medicated drinking and rebellious behavior. Takes approximately 5 minutes to complete.</p>	<p>Copyrighted</p> <p>Ordering information in Appendix B.</p>
<p>Adolescent Alcohol Involvement Scale (AAIS)</p> <p>Mayer, J. &amp; Filstead, W.J. (1979). The Adolescent Alcohol Involvement Scale: An instrument for measuring adolescents' use and misuse of alcohol. <i>J. of Studies on Alcohol</i>, 40, 291-300.</p>	<p>14-item screening questionnaire categorizes clients as: abstainers/infrequent drinkers, non-problem drinkers, alcohol misusers, and alcoholic-like drinkers. Takes approximately 5 minutes to complete.</p>	<p>Available in Appendix B.</p>
<p>Drug Use Screening Inventory (DUSI)</p> <p>Tarter, R.E. &amp; Hegedus, A.M. (1991). The Drug Use Screening Inventory: Its application in the valuation and treatment of alcohol and other drug abuse. <i>Alcohol Health and Research World</i>, 15, 65-75.</p>	<p>Written at a 5<sup>th</sup> grade reading level, 149 items screen for substance abuse, physical and mental problems, and psychosocial adjustment. Takes approximately 25 minutes to complete.</p>	<p>Ordering information in Appendix B.</p>
<p>Diagnostic Interview Schedule (DIS)</p> <p>Robins, L., Cottler, L., &amp; Keating, S. (1989). The NIMH diagnostic interview schedule. Version III, revised (DIS-III-R). Rockville, MD: National Institute on Mental Health.</p>	<p>Easily administered by non-clinical staff, this structured interview provides diagnostic classifications that correspond to DSM criteria. Takes approximately 15 to 20 minutes to complete.</p>	<p>Public domain.</p> <p>Ordering information in Appendix A.</p>

<b>Assessment Instruments For Use With Adolescents</b>		
<i>INSTRUMENT/ CITATION</i>	<i>DESCRIPTION</i>	<i>AVAILABILITY</i>
Rutgers Alcohol Problem Inventory (RAPI)  White, H.R. & Labouvie, E. W. (1989). Towards the assessment of adolescent problem drinking. <i>J. of Studies on Alcohol, 150</i> , 30–37.	Measure of alcohol problems developed for use with a student population.	Public domain.  Ordering information in Appendix B
Adolescent Drug Abuse Diagnosis (ADAD)  Friedman, A.S., & Utada, A. (1989). A method for diagnosing and planning the treatment of adolescent drug abusers: The Adolescent Drug Abuse Diagnosis (ADAD) instrument. <i>J. of Drug Education, 19</i> , 285–312.	150-item structured interview provides comprehensive assessment in 9 areas: medical, school, employment, social, family, psychological, legal, alcohol, drugs	Unpublished, available from author.  Ordering information in Appendix B.

## **LANGUAGE**

In selecting measurement instruments for use with any client, it is important to consider both the language fluency of the client and the language in which the test was developed. Clients whose native tongue is not English may be fluent enough in English to chat informally with you but may well miss, or be unable to express, important concepts in English. Particularly if the instrument requires the individual to read and write, fluency barriers can seriously distort the information you gather.

The translation of testing instruments is no simple matter. Tests that have been developed and validated in one language do not necessarily hold up when translated into another. Nuances that fail to translate along with the literal meaning of a word may cause important concepts to be missed by the individual being assessed. If you administer an English-language instrument to someone for whom English is a second language, be very careful in interpretation. Norms that have been developed to let you compare individuals with "normal "

groups are unlikely to be appropriate. For some tests, specific norms for special populations are available.

If you use a test that has been translated from one language to another, make sure that the translation has been validated with the population you intend to assess. For example, Spanish language versions may vary significantly for Mexican, Cuban, Puerto Rican, or Northern New Mexican peoples. If you are considering translating a test yourself (or having it done), do so carefully. A useful strategy in translating important material from one language to another is called "back translation." Quite simply this involves translating the material back and forth between the two languages a number of times, using different translators. If the back-translated version does not match the original version, this signals that a meaning has been contorted or lost. For test items where this is the case, your next step is to work with the translators to determine where the problem lies. It may be a language barrier insofar as the words do not directly translate and must be clarified for translation, or it may be a cultural difference in that the target language does not contain the words for the concept being expressed.

Interesting examples of the language versus culture differences are given by Richard Brislin<sup>18</sup>. He describes the process of translating a social-desirability scale from English into Chamorro, the native language on Guam, and an English marital relations questionnaire into Thai. One of the social desirability items was, "I like to gossip at times." Over the course of translating this item into Chamorro and back into English three times, the researchers learned that in Chamorro, there are separate words for male gossip and female gossip and no single translation was possible for the statement. After working out the meaning of the statement through three back-translations, the scientists and translators arrived at a satisfactory translation in the form of "Sometimes I like to talk about other people's business." In this case, the barrier was a linguistic one. In contrast, the Thai example illustrates a cultural barrier to translation. In this case, the sentence completion task used to study marital relations included the statement, "Sometimes a good quarrel is necessary because ...." Although disagreements are considered healthy aspects of marital relationships by the American population on which the questionnaire had been developed, spousal disagreement as a "good" thing was inconceivable to the Thai. Thus, an appropriate translation of this item was not possible.

Finally, with respect to language barriers to assessment instruments, consider whether the instruments depend on the client's ability to read. Illiteracy in the United States is

---

<sup>18</sup> Brislin, R. (1993). *Understanding Culture's Influence on Behavior*, New York: Harcourt Brace & Company, pp. 77-81.

growing, and reading problems may not be readily apparent from verbal interactions. If your assessment requires the individual to be able to read and write, confirm that s/he is able to do so prior to administering your tests. One easy way to do this is to ask the person to read aloud to you the instructions for one of the instruments.

## **CULTURE**

The cultural context in which an individual operates is not always readily apparent from simply knowing where he or she lives or what language he or she speaks. An individual, for instance, who is fluent in English and lives in the USA may nonetheless take his/her cues for appropriate behavior regarding substance use from a subculture with very different norms than the majority Anglo culture. Thus, assessment instruments that rely on clients' value judgments of usage patterns may entirely miss the mark if they are not sensitive to relevant cultural differences. As we saw in the Thai example above, it is not safe to assume that concepts commonly understood and agreed upon in one culture are necessarily to be found in another. For example, our research group at the University of New Mexico has been developing an instrument to measure attitudes about alcoholism. In doing so we have encountered substantial differences across locations and time in what people mean by "alcoholism." In other countries, we have found it difficult to use such an instrument at all because there is no social consensus about the meaning of "alcoholism," or use of the term is disapproved of as moralistic. For these reasons, it is important for you to review any instruments you intend to use with an eye toward the cultural assumptions they make when asking respondents to make judgments or offer opinions.

Cultural misinterpretations can occur not only in written tests you might have clients complete but also in therapist or interviewer assessments made during face-to-face interviews. For instance, reluctance to make eye contact is generally perceived by the majority Anglo culture to indicate shyness, or even dishonesty. However, among some cultures it is considered disrespectful for a subordinate (often a female) to make eye contact with an authority figure. Thus, a male Anglo interviewer assessing, for instance, a traditional Oriental woman might incorrectly interpret her lack of eye contact as a negative symptom rather than a positive sign of respect.

There are important differences in cultural norms regarding how much use of what substances constitutes "abuse," if indeed the concept exists at all. Instruments that rely exclusively on evaluative, general questions such as, "Do you drink too much?" will tell you about the respondent's

perception of his/her drinking status but may tell you nothing about how much he or she drinks. Depending upon the individual's personal norms and those that prevail within his or her culture, the threshold for judging "too much" drinking can vary considerably. In using assessment instruments that have been developed with a culture other than the one you are targeting, be wary of relying on questions that do not explicitly define the terms being used. For example, instead of "Do you drink too much?" it might be better, particularly in diverse populations, to ask for more specific information related to the particular standard you have in mind: "How many days during the past month did you drink any alcohol at all? What kind of alcohol beverages did you have to drink? And on days when you drank, how much did you drink?"

Patterns of substance use also differ from culture to culture and thus need to be considered in selecting your assessment instruments. Questions that ask about usage during a "typical week" may well miss important information about binge patterns. For example, a self-report of 12 drinks during a typical week might not capture much heavier drinking that occurs once every two or three weeks. Again, look at the questions you are asking clients and ask yourself whether they will capture the various usage patterns that may occur.

Finally, be cautious in assessing non-behavioral attributes such as "responsibility" or "relationship quality." Just as you run into cultural differences when you ask for judgments about appropriate behavior, asking about abstract (non-behavioral) things such as whether someone is "responsible" or has a "good" relationship leaves you with the problem of figuring out just what the answers mean. For instance, does a Cuban define responsibility the same way as a Puerto Rican, Japanese, or member of the Daughters of the American Revolution? Chances are there are substantial differences between their definitions and those differences will make it difficult for you to know exactly what your test results are telling you.

## **GENERAL RECOMMENDATIONS**

In light of the potential pitfalls of using assessment instruments for populations other than the ones with which they were developed, there are a number of general precautions we encourage you to observe.

- **Use objective measures.** Whenever possible, use assessment tools that rely on "objective" information that requires little or no interpretation. Such

specific information is most easily obtained from behavioral observations or reports. For example, frequency and quantity of substance use depend only on the individual to be able to count (or estimate) his/her usage. Similarly, asking for teacher or parent reports of *number of times* a child *did* something (*frequency* of a behavior) provides you with data from which you can draw your own conclusions instead of having to figure out whether "a lot" means the same thing to you and to the person being assessed. The following table illustrates the distinction between subjective and objective measures for some of the areas commonly assessed in treatment and prevention of alcohol abuse.

<b>Problem Area</b>	<b>Subjective Question</b>	<b>Objective Question</b>
Frequency of drinking	How regularly have you been drinking this month?  Very    Somewhat ...	How many days this month did you drink any alcohol?
Physical effects of alcohol	Does your use of alcohol affect your health?  Yes    No	Which of the following have you experienced during periods of heavy drinking?  <list relevant symptoms>
Life disruption	Is alcohol interfering with your life?  Yes    No	As a result of drinking, have you ever missed work?  <repeat for other relevant events>
General psychological status	Do you have any emotional problems beside those related to your drinking? Yes    No	Which of the following have you experienced?  <list symptoms of interest>

As you can see when you compare the sets of questions, the objective ones ask for more specific reports of behavior (how many days did you drink?) and experiences (have you experienced ...?). With this objective information, you can draw conclusions about the client's behavior relative to established criteria, rather than rely on their *opinions* and *impressions* of what is going on. Bear in mind we are not suggesting that clients' opinions and impressions are unimportant. To the contrary, they are very important. However, when it comes to program evaluation research with which

you want to be able to report about *groups* of clients, you need common measurement sticks, so to speak. Relying on individual expressions such as "I drink a lot" doesn't let you compare—because one person's "a lot" is not the same as another's.

- **Talk to the publisher or author.** Before making the decision to use an instrument for a special population, make sure it has been validated for that population. Talk to either the publisher or the author him- or herself and ask about the people with whom the test was developed. Describe the population you work with and ask whether the instrument will provide reliable and valid information when used in this context.

Unfortunately, not many instruments have been validated with minority populations. The following table lists the few that have. (Refer to Chapter 4 for descriptions.)

<b>Instruments Validated With Minority Populations</b>	
<b>INSTRUMENT</b>	<b>COMMENTS</b>
Alcohol Use Disorders Identification Test (AUDIT)	The AUDIT has been validated by the World Health Organization on a large cross-cultural sample.  Available in Appendix A.
Diagnostic Interview Schedule (DIS)	This interview form has been translated into Spanish and normed accordingly. There is also a version available for use with children and adolescents.  Ordering information in Appendix A.
Drinkers Inventory of Consequences (DrInC)  Versions 2-L and 2-R	Translated into Spanish. Hispanics responded similarly to non-Hispanics on the English language version.  English version in Appendix A.  Spanish version in Appendix B.
Goal Attainment Scale	Available in Appendix A.

## **SUMMARY**

Whenever you assess clients who differ culturally, linguistically, or in age from the population with which the assessment instrument was built, you risk getting misleading results. In selecting your instruments, look for ones that

gather objective information that can potentially be verified and that included a representative sample of your special population in their development.

Now return to Worksheet #9 and review your measurement tool selections to ensure they are indeed appropriate for your population. If you are not sure of them, consult with an expert.

## **6—FOLLOW-UP EVALUATION**

We include a whole chapter to discuss program follow-up because it is a vital component of evaluation for substance abuse prevention and treatment, yet is overlooked by many programs. Too often, due to budget constraints and misconceptions, programs keep excellent track of clients while they are actively involved in the program but fail to collect information about what happens when they leave. Even programs that provide extended maintenance (aftercare) contact often fail to systematically keep track of clients at this stage and beyond. (In plain words, they don't know if the program sticks.)

Substance abuse programs, indeed all behavior change programs, are intended to result in lasting changes in the people they touch. Whether you are implementing a drug prevention program in elementary schools or treating long-time heroin addicts in a residential program, your objective is to trigger changes that last beyond the time you actively work with clients. So, you need to know how your intervention affects clients while you actively work with them and you need to know how your intervention affects them after you work with them. Follow-up tells you what happens *after*.

### **WHAT IS FOLLOW-UP EVALUATION?**

Follow-up is simply that stage of your program evaluation project that occurs when you are no longer actively working with clients. Include it in your program evaluation project design right from the start.

If your program routinely offers extended contact (aftercare, maintenance sessions) to clients, be careful not to prematurely call it follow-up evaluation. Those contacts provide follow-up data only if they systematically gather information about what clients are doing between contacts. You can, in many cases, use the same client sampling strategies, instruments, and data collection methods for follow-up as you did for the first half of your project. In fact, the single biggest difference between the first data gathering phase and the follow-up will be, unfortunately, how much more difficult it is to get to clients after they have "graduated" from your program. However, *it is possible* and the value added to your project and your program makes the effort well worth while.

## **WHY DO FOLLOW-UP?**

We've already hinted at the value of including follow-up as a routine part of your program but the benefits to both client and program are worth the emphasis provided by repetition.

- The information gained about which program effects "stick" and which don't is valuable in helping you to further refine the program. Without this information, revisions made to improve the program's long-term results remain pretty much shots in the dark.
- As a practitioner, the feedback you get from clients in the months after you are "done" with them can provide critical insight into how your own practices can be improved. Without this feedback, you cannot be sure that what you are doing has the intended effect. In addition, it is only with this level of feedback that you can begin to identify which approaches work best with which clients.
- Follow-up data can be a powerful ally when it comes time to justify your professional existence. Without systematic follow-up, you must rely on anecdotal reports and after-the-fact surveys to document the program's effectiveness. Increasingly, this level of data is not acceptable to accrediting and funding agencies. By integrating follow-up into your overall program plan, you will always be ready with needed, reliable information.
- Continued contact with clients can help avert relapse, or nip it in the bud. Follow-up sessions, even those designed primarily to only gather information, can work like booster sessions to reinforce changes made during the active phase of the program.
- Follow-up contacts with prevention program graduates can also serve to reinforce lessons learned and changes made during the earlier phase of the program.
- Information that has been systematically gathered over the course of the program and beyond can be easily summarized and shared with other programs. By sharing information in this way, facilities can minimize the duplication of costly exploratory efforts and more easily learn from one another.

As you can see, building follow-up procedures into your program, whether you offer prevention or treatment or both, can only improve the quality of your service. Furthermore, failing to gather this information can inadvertently harm the program by overlooking needed changes and blindly implementing unneeded ones.

## **HOW DO YOU DO FOLLOW-UP?**

There are two parts to the "how" question. One is "What tools should you use?" and the other is, "How should they be administered?" Tools were discussed at length in Chapters 4 and 5 so we will not repeat ourselves here. However, we will remind you to carefully consider what it is you want to know and make sure the instrument you select provides that information. When possible, it is a good idea to use the same instrument(s) during the program and for follow-up evaluation. That way you can look at the change in clients' scores to determine whether improvement has occurred or been maintained.

There are essentially three ways you can gather your follow-up information. You can assign a staff member to meet with clients face-to-face, speak to clients over the telephone, or contact them by mail. Each method has pros and cons that you should consider in your selection.

<b>Face-To-Face Follow-Up</b>	
Pros	<ul style="list-style-type: none"> <li>• Conveys a personal interest in the client thereby increasing the likelihood s/he will remain in contact with the program.</li> <li>• Allows for flexibility in pursuing unexpected topics that may come up during the course of the interaction.</li> <li>• Provides non-verbal information such as client behavior, appearance, etc.</li> <li>• Allows you to gather biomedical evidence of substance use or nonuse.</li> </ul>
Cons	<ul style="list-style-type: none"> <li>• The challenge of scheduling mutually agreeable times may reduce the number of follow-up participants.</li> <li>• Interview locations that are not convenient for the client will reduce the likelihood of attendance.</li> </ul>

<b>Telephone Follow-Up</b>	
Pros	<ul style="list-style-type: none"><li>• Easier to schedule phone appointments since no travel is involved.</li><li>• Allows for flexibility in pursuing unexpected topics that arise in the course of the conversation.</li></ul>
Cons	<ul style="list-style-type: none"><li>• Limits your contacts to only those clients who have telephones.</li><li>• Important nonverbal messages may be missed.</li><li>• Less nonverbal information to verify and encourage honesty.</li><li>• Cannot obtain biomedical evidence of usage or nonuse.</li></ul>

<b>Mail Follow-Up</b>	
Pros	<ul style="list-style-type: none"> <li>• Eliminates the scheduling task.</li> <li>• Places the least perceived demand on the client to “please” the interviewer since the client’s only interaction will be with printed material.</li> <li>• Is the least expensive means of gathering data (no interview personnel costs, telephone charges)</li> </ul>
Cons	<ul style="list-style-type: none"> <li>• Questionnaires are apt to be lost, forgotten, or otherwise not find their way back to you.</li> <li>• Does not provide the same degree of personalization as face-to-face and telephone contacts do. This may not accomplish the objective of making clients feel like they are still “in” the program.</li> <li>• Nonverbal messages are lost.</li> <li>• Allows no flexibility in the nature of the questions.</li> <li>• Relies on client’s reading ability.</li> </ul>

As you can see, there is no one best way to collect the information you need; every method has its pros and cons. The “best” method will be the one that most closely matches the needs of your organization, the available personnel and budget, and the nature of the information you are seeking.

Keep in mind that you need not limit yourself to only one approach. Begin with one system and then try others as the need arises. For instance, you might decide that face-to-face follow-up is the way to go for your needs but some of your clients simply cannot return for these meetings. For those clients, telephone or mail follow-up can be used.

### **WHO SHOULD DO FOLLOW-UP AND WHOM DO YOU QUESTION?**

Depending upon the nature of your program, the follow-up information you gather will vary and so the appropriate personnel to gather this information will also vary. The following discussion will help you figure out the best strategy for your setting.

**Interview by the staff member originally involved with the client.** There are both pros and cons to having the follow-up interview conducted by the individual’s practitioner, teacher, or case worker. Obviously, on the pro side there may be an established rapport with the original care-giver and his/her familiarity with the case. There is also the

advantage for the staff member of gaining first-hand feedback about his/her impact on the client. However, on the negative side we have two possibilities. One is that the client may be feel the need to please the interviewer by reporting better outcomes than really occurred. The other is that the interviewer, having originally been the client's primary contact, may have a personal bias and "hear" the client's reports more positively or negatively than they actually are.

**Interview by a staff member who has not had prior contact with the client.** In this case we do not have the problem of clients feeling uncomfortable reporting "bad news" about themselves or negative perceptions of the program to their original care-giver. Furthermore, the "stranger" interviewer will not be as likely to filter his/her perceptions of the client through the bias of a history of having worked together. On the negative side, being interviewed by a stranger does place the client in the position of having to open up to someone s/he may not feel at ease with.

**Client completes a questionnaire without interaction with a staff member.** Simply providing clients with a printed questionnaire and asking them to complete it eliminates the concerns about distorted results due to client-staff interactions. This method is also considerably less expensive because it incurs no additional personnel costs. (A receptionist can hand questionnaires to clients or questionnaires may be mailed to clients' homes.) An additional benefit of client questionnaires is that if you use a multiple choice format, you avoid the problem of deciphering people's handwriting. On the down side, it allows for no flexibility in addressing or expanding on unexpected issues, nor the ability to clarify questions for clients. Finally, if a client's reading ability is limited, so too will be his or her ability to complete the questionnaire.

**Information from collaterals.** Whether clients complete questionnaires on their own or are interviewed by staff members, there is always the concern they may provide inaccurate data. One way to verify the accuracy of self-report is to obtain permission from clients to contact one or two people in their circle who know them well. Ideally, these "collaterals" will be people who are reasonably reliable and have no vested interest in misrepresenting the client's behavior to you. You can then simply ask the collateral parallel questions to those asked of the client and compare them. Of course, contact with a client's social system may be clinically useful as well.

## **WHEN TO DO FOLLOW-UP?**

In most cases, you will want your follow-up activities to serve two purposes. One is to improve your service to clients after they have completed the initial change phase of the program and the other is to provide data you can use for program evaluation.

To best assist clients in the months following treatment completion, follow-up contacts might be planned for the times clients are most likely to experience difficulty. The research on substance abuse treatment outcomes indicates that relapse is most likely to occur within six months of program completion. Thus, follow-up contact within the first six months is optimal and most programs plan a 3- and/or 6-month post-treatment follow-up. However, such short-term outcome does not necessarily predict long-term outcome. Thus, for program evaluation purposes, longer-term contacts are in order and, we might add, sometimes required by agencies that evaluate programs. Thus, to meet both the need for aftercare and evaluation data, good contact points are 3, 6, and/or 12 months.

If you are planning follow-up for a prevention program, "contacts" as discussed above will most likely occur in the form of participant surveys in the time following program implementation. For example, if you were setting up a gang-resistance curriculum for all sixth graders in a school district, you might survey attitudes toward and involvement with gangs among sixth graders and seventh graders prior to implementing the program and then survey seventh graders the following year (your ex-sixth graders). Obviously you would be looking for a holding pattern or reduction in pro-gang attitude or involvement from sixth to seventh grade among the same group. You would also expect less involvement among the seventh graders with whom you had worked in sixth grade than among the original (untouched) seventh graders. In terms of timing the follow-up surveys, it would probably be good to plan multiple contacts during the second year of the study, if possible, to see whether the program effects change over time.

Ultimately, the timing of follow-up contacts is best determined by careful consideration of the client population you serve, the natural evolution of the disorders you target, and the resources accessible to your organization.

## **A REMINDER ABOUT FOLLOW-UP MEASUREMENT INSTRUMENTS**

In selecting the measurement instruments to use for follow-up, give careful consideration to the timespan addressed by the questions. Many instruments, particularly those that measure substance abuse, phrase their questions either in terms of lifetime usage or specify an inappropriate time frame for your study. When you want to identify *change* in behavior over the relatively short duration of a treatment program, you need instruments that ask about usage (or symptomology or distress, etc.) in the more recent past.

For example, assume you offer a 12 week treatment program with follow-up contacts at 3-, 6-, and 12-months after treatment. During follow-up you will probably want to ask about client behavior for the 3 months prior to each contact rather than a longer period that would confuse behavior during and before treatment with that during the different segments of post-treatment time you are interested in. Always *examine* the questions to be sure they are appropriate for follow-up.

## **INVOLVING CLIENTS IN FOLLOW-UP**

To successfully implement follow-up in your program, you will naturally need to involve clients so that they show up (in person or on paper) for your follow-up contacts. In spite of the fact that these contacts may be "good for them," it will not be easy to obtain follow-up data on everyone you target. However, there are steps you can take to maximize the number of clients who cooperate.

- At intake, present the program as a "whole" that includes two phases: initial active involvement and follow-up. If clients expect to have continued contact with you post-treatment, the task of making that contact will be significantly simplified.
- Give careful consideration to the nature of your clientele. If you serve a middle class or affluent community in which transportation to your facility is readily accessible (along with baby-sitters, cold-weather clothing, and the like), follow-up sessions at the facility may be ideal. On the other hand, if your population is more spread out, has poor access to reliable transportation, and so on, telephone or mail contacts may provide the more reasonable approach.
- Consider the feasibility of requiring from clients at the outset of the program a monetary deposit that would be refunded upon completion of the "whole" program. The refund serves as an incentive to return for follow-up.
- Assure clients that whether or not their treatment gains are maintained throughout the follow-up period, the information they provide is helpful both for the purpose of improving the program in general and also in helping program personnel know when to reach out to the individual client and intervene in a downward spiral.

Whether your primary objective in conducting follow-up is to obtain program outcome data, to justify further funding, or to minimize relapse in clients, these data provide the essential building blocks for strengthening your program. Thus, in all cases, your clients ultimately benefit.

## **PLAN YOUR FOLLOW-UP**

Use Worksheet #10 on page 109 to plan your follow-up assessments. Keep your evaluation question firmly in mind and refer to the notes you made in Worksheets #8 and #9. Consider the best way to gather the post-program information you might need and the best time to do it. If your follow-up plans make it necessary for you to revise your measurement tool selection, now is the time to do so.

This page intentionally left blank.

WORKSHEET #10. PLAN FOLLOW-UP ASSESSMENTS.

*What method(s) will you use? (Face-to-face, telephone, mail)*

*Who will conduct the follow-up contacts?*

*When will contacts occur?*

*Are the measurement instruments selected in Worksheet #9 appropriate for follow-up?*

This page intentionally left blank.

## **SUMMARY**

In designing your follow-up procedures, or reviewing those already in place, give careful thought to the population you serve and the information you most need. Follow-up strategies that work well in one setting with one client population may not be completely appropriate in another. Design your approach to maximize the quality of the information you gather—but do so in a fashion that is most likely to result in clients' cooperation. Otherwise, you may end up with an elegant research design, but no data!.

It has been said that doing treatment without follow-up makes as much sense as playing golf in the fog<sup>19</sup>. You swing at the ball (treat the client or offer prevention services) but then it disappears into the fog and you have no idea where it lands. In other words, without follow-up, once clients walk out of your door, you've no way of knowing how well your intervention is guiding the client forward in a positive direction. Sadly, the price of this blind swing is paid three times over. It is paid by the practitioner who is robbed of seeing his/her work fulfilled, by the program that lacks the feedback needed to grow, and also by the client who fails to receive the ongoing support that is part of follow-up.

---

<sup>19</sup> Ziskin, J. (1970). Coping with psychiatric and psychological testimony. Beverly Hills, CA: Law and Psychology Press.

This page intentionally left blank.

## **7—MAKING SENSE OF YOUR DATA**

Having decided on your research question, designed your study, gotten everyone on board, selected your instruments, and identified your participants, you are ready to begin. Right? Wrong.

Before you begin gathering data, give some thought to how you will make sense of those data. Like everything else we have talked about in this manual, planning ahead for how you will handle data makes the difference between a simple, clean operation and, quite frankly, a mess.

If you have ever taken a research design or statistics course, you know there a million and one ways to organize, analyze, and present data. You could, if you wish, get quite fancy with your results. Or, you could simply apply some basic organization and elementary arithmetic skills to summarize your findings into a meaningful story about your project. For the purposes of this chapter, we will assume you wish to do the latter.

The remainder of this chapter presents five sections that could be thought of as steps. If you take the time to carefully work through each one, you will be able to extract a meaningful story from your data and your project at its conclusion.

### **PROTECT CLIENT CONFIDENTIALITY**

Remember that no matter how beneficial program evaluation may be for clients in the short or long run, nobody wants to feel like a guinea pig and no one wants to read their personal information on the company bulletin board. In other words, protecting client confidentiality is the first step in planning how you will gather data. In fact, it is a most serious professional responsibility.

There are a number of ways you can keep clients' personal information secure and confidential through your study.

- Keep all records in locked files, and password protect any computer files.
- Assign each client a unique participant number. Keep a master list connecting clients' names to their numbers under lock and key. At the start of the study, place all the assessment instruments you plan to use for each client in his/her file with the appropriate number marked on

them. At the conclusion of the study, separate the numbered instruments you will analyze from their identifying files. You can match up pre- and post-treatment instruments by their numbers but no one looking at your data will be able to tell which clients the data came from. Of course, if you run into a problem and need to identify the client for a particular instrument, you have your master list to refer to.

- Let participants remain anonymous. Instruct clients not to put their names on anything. (This strategy only works if you do not need to match up multiple measurements from the same clients.)
- If you are primarily using material that is or needs to become a permanent part of clients' files, the confidentiality restrictions that apply to the practitioner/client relationship remain in effect. However, you should assign participant numbers to use when you begin transcribing data from client files to your data spreadsheet.

## **CODE WORDY INFORMATION**

If you have chosen to use any of the questionnaires we recommend in Chapter 4, half this step is done for you. All of those instruments provide numeric scores that you can easily summarize and compare. This is true, as well, if you have designed your own questionnaires in multiple choice format so that you can easily count how many people selected one option versus another. However, there is likely to be other client information you have collected that is in verbal form (e.g., presenting problem, gender). Some of it, like gender, is easy to summarize as it is (e.g., 34 males and 22 females participated in this study). Other information, however, does not lend itself to counting as easily.

Unless you have very few participants in your study, you will need a way to summarize the verbal data from groups of clients into meaningful "chunks." The simplest way we know of to do that is to code it. For instance, let's say you are investigating whether clients' reasons for seeking treatment have an effect on whether or not they complete your program. To test this, you have asked 30 alcohol treatment clients what brought them to your program

### ***Reasons for Seeking Treatment Hodgepodge List***

DWI court referral  
Worried about my drinking  
Spouse threatened divorce  
Employee Assistance Program referral  
Family forced me here  
Cracked up my car  
Kids embarrassed by my drinking  
Marriage on the rocks  
etc.

and you have a hodgepodge list of their answers. It reads something like the list in the example box. Obviously, you can hardly predict anything from 30 idiosyncratic ways of speaking. So, do the following.

Review the list for items that can be grouped together. For instance, you might group "DWI court referral" and "cracked up my car" under the heading "DWI." Similarly, "spouse threatened ..," "family forced..," "kids embarrassed ..," and "marriage on .." could be coded as "Family." Presumably there would be other EAP referrals among the 30 clients and so "EAP" would be a reasonable category under which to gather all of these. Depending upon the nature of your data, of course, the categories will vary. Keep in mind as well that there will almost always be a few items that just don't seem to fit with any others. Classify them under that old catchall, "Misc."

Once you have coded each item with a category code, it is easy to summarize all the responses into a few meaningful chunks that will be easy to work with and report. If you look at our example here, you can see we are now able to easily identify a group of clients who entered treatment because of DWI problems, another who were pressed to enroll by their families, and so on. With these groups identified, we can now begin to compare their progress through the program.

<b>Reasons for Seeking Treatment Meaningfully Coded</b>	
DWI	2
Family	4
EAP	1
Misc.	1

The same strategy can be applied to information you gather orally or from open-ended written questions. Review your interview notes or the written responses and look for common themes or categories in which to group them. The time required by this process, though, will illustrate why we advised you to keep open-ended questions to a minimum.

## **PREPARE & ORGANIZE YOUR DATA**

If you have a statistician at your disposal, this step is the easiest of all. Simply do the following:

1. Explain to the statistician what your research questions are.
2. Tell her/him how you are testing them.
3. Decide together what format your data should take.
4. Gather your data.
5. Hand over all your data.
6. Go on vacation until s/he informs you that the data are analyzed.

On the other hand, if you are not blessed with an in-house stats whiz, you will probably want to get your hands on a good computer spreadsheet. There are a number of excellent packages available for under \$400. Before you purchase anything, though, check with your accounting department. They may already own a spreadsheet that will serve your purposes nicely.

If your facility is not equipped with computers or your budget does not allow for spreadsheets, don't despair. Datasets were managed long before PCs came to rule the workplace. Data can be managed by hand. It just takes a little longer.

We will begin by showing you how to organize your dataset by hand and then explain how to enter it on a spreadsheet. Unless you are already familiar with computer spreadsheets, you should read the *by-hand* instructions even if you plan to use a spreadsheet. The basic organization is the same and understanding the manual set-up will make the computer set-up much easier.

**Setting up your data by hand.** The first thing to do when organizing your data is remember that you will want to summarize like pieces of information. For example, you will probably want to quickly count the number of males and females or scan down a list to see how many sessions each participant attended. Obviously, if you have the data from each client in a separate folder, organized in a different order, skimming, scanning, and otherwise summarizing will be a nightmare. Here's what to do.

1. Create a spreadsheet (the old fashioned paper kind) on which to record the data you gather from your study. If you don't have fancy lined and columned paper, use simple lined paper and draw as many columns on it as you have pieces of information about each study participant. (Steps 2 and 3 below will help you figure out how many columns to draw.) Spreadsheet paper with ready-made rows and columns can be found at many office supply stores.
2. Decide on a meaningful order in which to record the data in the spreadsheet's columns. For instance, if you administered the QFV-30 at pre-treatment, post-treatment, and again at the 3 month follow-up, you will probably want to be able to quickly scan your data and see what kind of changes clients exhibited across the three administrations. Thus, it makes sense to record the three scores side by side in the order they were taken.

3. Next, try to anticipate any *derived scores* you may want to examine. Derived scores are simply scores you calculate from actual gathered data. For example, if you are comparing two treatment groups, you may want to know whether there was a larger decrease in alcohol related problems in one group than the other. Thus, you might have pre- and post-treatment DrInC-2R scores for each group and calculate a change score for each client. That way, you can examine the change scores for each group and decide whether one showed larger decreases than the other.
4. If you will use derived scores, plan to position them on the spreadsheet close to the gathered data they are based on. There are two reasons to do this. One is it will be easier to calculate them if you do not have to go to one section of the spreadsheet to look up the numbers and another to record the calculations. The other reason is that it simplifies the task of finding meaningful relationships among your data if related pieces of information are close to one another.
5. Label each column on your spreadsheet with the data that will be recorded in it. Don't forget to include columns for subject identification numbers.
6. Record your data on the spreadsheet by using one row for each client/subject and recording his/her data in its corresponding columns. If you have coded wordy data, record your short codes rather than trying to transcribe the wordy raw data. See the example that follows.

Subj. #	Sex	Reason for seeking Tx	Pre-Tx. DrInC-2R	Post-Tx. DrInC-2R	Change in DrInC-2R
1	F	DWI	60	42	-18
2	M	FAMILY	74	67	-7
3	M	DWI	69	60	-9
4	M	MISC.	58	65	+7

Setting up data on a computer spreadsheet. Organizing your data on an electronic spreadsheet is essentially the same as that by hand. In fact, the only real difference is that instead of recording the information on paper, you do so at the computer. Most<sup>20</sup> electronic spreadsheets are designed so

<sup>20</sup> We're tempted to say "all" electronic spreadsheets but there is someone in every crowd who owns a piece of software that doesn't fit the

that you enter your data just as you would on a paper spreadsheet. When you fire up the program, it presents a grid of rows and columns on which you organize and enter your data exactly as we instructed above. The only difference we would recommend is that you keep the CAPS LOCK key on if you are entering any letters. This avoids the problem of having the same information coded in both upper and lower case. When that happens, the computer may miss entries when searching.

There are advantages to using an electronic spreadsheet. The most obvious is that it can very quickly do all the counting and calculating that you need done—and do it flawlessly. Another advantage is that if you need to re-arrange your columns or insert a client row between rows already entered, you can do this without the arduous erasing and re-copying that paper and pencil requires.

Take some time now to list the variables to include on your spreadsheet. In addition to the main outcomes you are interested in tracking, remember to list the other pieces of client and program information that you may need to consider as you look at outcomes; information such as client number, age, gender, therapist, initial diagnosis, etc. Use Worksheet #11 on page 120 to do this.

---

mold. Chances are, however, that if your spreadsheet was purchased within the past few years, what we say here is true for your software. The three most widely used electronic spreadsheets are Microsoft Excel, Lotus 1-2-3, and Quattro-Pro.

This page intentionally left blank.

WORKSHEET #11. DECIDE WHAT YOU WILL ANALYZE.

LIST EACH VARIABLE THAT WILL REQUIRE A COLUMN ON YOUR SPREADSHEET.

Short Variable Name  
(Column Header)

Variable Description

This page intentionally left blank.

## **CHECK YOUR DATA**

Whether you use a computer or manage your data by hand, it is always important to check the accuracy with which the data were recorded. In a perfect world, or with a reasonably small dataset, you can proofread each entry against its original source. However, if you have a fairly large dataset, and do not live in a perfect world, you may need to take some short-cuts in checking your data. There are a few ways to do this.

- **Conduct a random check.** Carefully proof read a random handful of entries for clients/rows at the beginning, middle, and end of your spreadsheet. Aside from catching entry errors for these particular clients, this will also alert you to any pattern of mis-entries you may have fallen into as you entered your data. For example, you may inadvertently have begun writing or typing "B" instead of "M" for Male at some point in the process. Although "B" (for boy) might make semantic sense, when you are rapidly skimming the column looking for M's, or the computer is counting them, the boys will not be counted with the men!
- **Check for ridiculous entries.** Under each column heading, you expect a certain range or type of information to be entered. For instance, if your study involved only clients under 18 years of age, you would expect all entries in the age column to be less than 18. Further, if you knew that your youngest participant was 13, all entries should be from 13 to 17. By asking the computer to give you the range of values for that column, or by scanning it yourself, you can check for values that fall outside the expected range. In the age example just given, a range of 13 to anything over 17 would obviously indicate an entry error.
- **Do some calculations.** If you are using a computer and have a reasonably good idea as to what the average value of the column should be, have the machine generate a mean for you. If the mean value is way off from where you expect it to be, do a more careful check of that column. Chances are there a number of mis-entries. Using the age example again, let's say you know that most of the kids in the study were 13 years old with only a handful being 14 through 17. You would expect the average age in this case to be around 13. If the calculated average comes out much higher than that, you can bet you mis-entered a number of ages higher than they should be. Go back and check the individual entries.

## **ANALYZE YOUR DATA**

The most exciting part of any evaluation project is analyzing the data. It is in this step that you get to see what fruits your labor has borne. Data analysis, like everything else we've talked about in this manual can be simple or highly complex—depending on how you approach it. If you have clearly defined your research question(s) and designed your project carefully, you can get answers fairly easily.

When analyzing a dataset, there are basically two main tasks you want to perform. One is to simply describe the participants (e.g., how many males/females, how long did they participate, how many completed which assessment step, what scores were obtained, etc.). The other is to evaluate the results obtained.

If you are working by hand and really enjoy doing manual calculations, all you will need is a pencil. If you are working by hand and are not particularly fond of manual calculations, we recommend a calculator. If you are using an electronic spreadsheet, you already have all the tools you need.

### **Describe the Dataset**

You can adequately describe your dataset and the people from whom it was generated by reporting three things about each relevant variable: the range, mean, and standard deviation.

**Range:** The lowest and highest number in a set of numbers.

**Mean:** The average of a set of numbers.

**Standard Deviation:** The average amount that the numbers in a set deviate from the set's mean.

The *range* of a set of numbers is simply the lowest and highest numbers in the set. If you are working by hand, simply scan the column in question and find the lowest and highest numbers.

If you use an electronic spreadsheet, you can quickly

find the range of a column by doing the following:

1. Save your spreadsheet.
2. Select the column in question and have the program sort it in ascending order.
3. Jot down the first number in the column and the last one.
4. Go to the edit menu and undo your last action. (**DO NOT FORGET** to undo this action. If you do, the values for

this variable will no longer match the clients who generated them!)

The *mean* is the average of a set of a numbers. Calculating the mean lets you describe the whole set with just one number. For instance, imagine you had 100 people rate on a scale of 1 to 7 how helpful they found a new smoking cessation pamphlet. You could try to report their opinions by listing all 100 ratings—an onerous task at best. Or, you could report the range (see above) and the mean, or average. With these two measures, you can describe all 100 responses fairly well. You would know the highest and lowest rating as well as the one that is most representative of the group.

Calculating the mean is a breeze if you are using an electronic spreadsheet. Most will let you either type in a command that generates a mean or select mean or average from a data analysis menu. For example, with Microsoft's Excel, you can quickly ask for a mean by typing =AVERAGE (CELL#.CELL#) where CELL# are the coordinates of the first and last cells in the column.

If you need to calculate means by hand (or with a calculator), here's what you do. Add up all the numbers in the set and divide the total by the number of figures you added. Simple.

If you know the range and mean of a number set, you already know quite a bit about the set. The range tells you how high and low it goes and the mean tells you what the most typical value is. However, there is one more thing you should know about the set to really have a sense for what it looks like. That is, how spread out or clumped together are the numbers?

Consider the numbers in our Calculate a Mean example. If you look at the eight numbers themselves and the mean, you can feel pretty good that the mean describes the average value. However, look at the numbers in the next mean example. Here

**Calculate a Mean**

1. Count how many numbers are in your set.
2. Add them all up.
3. Divide the total by the number of numbers in the set.

*Example.*

The set: 3 6 2 4 8 9 7 6  
There are 8 numbers in the set.  
 $3+6+2+4+8+9+7+6 = 46$   
 $46 / 8 = 5.75 =$  the mean of the set.

**The Next Mean Example**

The set: 1 1 1 1 1 1 1 39  
There are 8 numbers in the set.  
 $1+1+1+1+1+1+1+39 = 46$   
 $46 / 8 = 5.75$

we have eight numbers again that average to the same 5.75 mean. Yet look at them. If you had only the range and mean to go on, you would not have a very clear sense of what this dataset really looks like. The missing piece in this puzzle is the *standard deviation*. It tells you the average distance the members of the set are from the set's mean.

### **Standard Deviation Calculation**

1. Calculate the mean of the set.
2. Subtract each number from the mean and square the answer.
3. Add up all the squared answers.
4. Divide the sum by the number of numbers there are in the set less 1.
5. Calculate the square root of the answer to Step 4.

The standard deviation can be easily generated by an electronic spreadsheet or, with a little more effort, calculated by hand.

The standard deviation for our first mean example was 2.45, indicating that the average dispersion of the set's members was only 2.45 units from the mean. On the other hand, when we calculate the standard deviation for the second sample data set, we get 13.43. This much larger number tells us the data are fairly spread out around their mean.

### **Standard Deviation Example**

The set: 1 1 1 1 1 1 1 39  
The mean: 5.75  
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $1 - 5.75 = -4.75^2 = 22.56$   
 $39 - 5.75 = 33.25^2 = 1105.56$   
Sum = 1263.48

$$1263.48 / 7 = 180.49$$

$$\text{Square root of } 180.49 = 13.43 = \text{Standard Deviation}$$

### **Evaluate the Results**

Having described your dataset, you are ready to evaluate it. Depending upon the nature of your project design, evaluation will take one of two forms. It will either consist of looking at the descriptive statistics you just calculated (i.e., the range, mean, and standard deviation) and deciding whether they provide any meaningful answers about your research question. Or, it will call for statistical analysis.

There are quite a few evaluations you can make without using any more statistics than the three we have already discussed. You can:

☞ Determine whether average assessment scores have noticeably changed over time or differ largely between groups.

☞ Report whether there was a greater incidence (frequency) of something occurring between one time and another or between groups.

☞ Identify the presence or absence of something you wanted to eliminate or trigger.

☞ Track increases or decreases in client traffic.

☞ Identify headway made in extending agency resources to cover community needs.

One thing statistics allow you to do over and above what good judgment enables, is conclude whether the differences in means between groups is more than just luck (or bad luck as the case may be). With statistical tests, we can conclude that the differences we observe are *significant*. In other words, there is only a very slight chance that they are the result of chance. These tests are called *tests of significance*. They add credibility to the conclusions you draw about your results.

Tests of significance typically require a fair bit of calculation so if you are working your data by hand, you may not want to undertake them. If you are using an electronic spreadsheet or other computer software to manage your data, it takes only a moment or two to generate the test results.

There are two tests of significance that you are most likely to want to use and both are available on most electronic spreadsheets.

#### THE T-TEST

The *t-test* tells you whether the difference you observe between the means of two groups is a *statistically significant* one or if it was most likely due to chance. When you run a *t-test* on two sets of numbers, you get a value for *t* and a probability of having gotten that value by chance.

We conclude that a difference is statistically significant if the probability of obtaining the *t-statistic* we did is 5% or

less. Your spreadsheet will give you the t-statistic and the probability value expressed as "p."

For example, consider the fictional data below. They are arranged just as they would be on an electronic spreadsheet. In column A we have clients' ratings of the strength of their urges to drink during the second week of treatment without training in urge management and, in column B, strength ratings by clients who were trained. The means of the two groups are 13.7 and 8.6 respectively. The question is: Do the means reflect great enough differences in the two sets of data for us to conclude the difference is *most likely* due to the differential training?

	A	B
1	W/o training	W/training
2	12	8
3	20	5
4	18	10
5	3	18
6	15	4
7	10	5
8	17	6
9	18	12
10	16	7
11	8	11

To answer this question, we select the *t*-test option on our spreadsheet and tell it the two sets of numbers to compare are located in cells A2..A11 and B2..B11. (The '..' means "through.") In seconds, the spreadsheet comes back with the test results. Most spreadsheets will probably give you more information than you need. Don't let the unfamiliar material throw you. You are looking for the *t*-statistic and the probability of *t*. Microsoft Excel reports these as *t Stat* and *P(T<=t) one tail* and *P(T<=t) two tail*. For most purposes, you will be interested in the two tail probability. If that value is .05 or smaller, you conclude the means are *statistically different*, or *significant*.

When we ran the dataset from the fictional urge training study through Microsoft Excel, we got a *t* value of 2.36 and a two tail probability of .03. We concluded that the difference observed was probably not due to chance and likely the result of the training intervention.

#### ANALYSIS OF VARIANCE

If you need to compare more than two groups, the *t*-test will not help you. Instead, you will need to use a related test called *Analysis of Variance*, or *ANOVA* for short. The general principle behind it is similar to the *t*-test. However, instead of telling you whether only two means differ, it tells you whether the data in 2 or more groups differ significantly. *ANOVA* is also available on most electronic spreadsheets. When you run it, you are interested in obtaining the *F statistic* (instead of the *t* as above) and looking for a probability value of .05 or smaller.

#### PEARSON CORRELATION

The final statistic we will talk about is the Pearson correlation. A correlation simply tells you whether two sets

of numbers tend to go up or go down together, opposite from one another, or have no relationship to speak of. The correlation statistic we calculate, called the Pearson correlation coefficient, is the *r*-statistic. The *r*-statistic has a possible range from -1 to +1. If two sets of number follow each other perfectly in the same direction, go up or down in tandem, they have an  $r=+1$ . If they go perfectly in the opposite direction, their  $r=-1$ . If you get an  $r=0$ , that means the two variables do not correlate at all. That is, the size of one has no relationship to the size of the other.

As always, things are easier with an example. Let's say you are interested in knowing whether there is any fact behind therapists' casual observations that the older the client, the longer s/he tends to stay in treatment. You might gather client ages and their length of stay into columns like this.

Client #	Age in years	Stay in weeks
1	18	3
2	21	5
3	34	10
4	19	5
5	25	8
6	60	18
7	52	17
8	46	16
9	28	12
10	19	10
11	19	5
12	52	18
13	54	19
14	58	20
15	47	17
16	49	20
17	36	15
18	39	16
19	42	18
20	28	15

We entered the age and stay data into Excel in two columns and got the following print-out.

	Column 1 (Age)	Column 2 (Stay)
Column 1 (Age)	1	
Column 2 (Stay)	0.905114	1

Recall that a perfect positive correlation is 1 and would mean that, without exception, the older the client the longer the stay. When we correlated Age with Stay, we came up with a correlation of  $r=0.905114$  which tells us there is a very

strong relationship between age and stay, albeit not a perfect relationship. Had the value of  $r$  been negative, it would have meant that the older the client, the shorter the stay.

Correlation analysis is useful when you want to investigate the relationship between two variables that you suspect "travel" together. It does not tell you anything about the differences between the effects of your independent variables.

Although most statistical analysis software provides levels of significant for this test, that is,  $p$  values, most spreadsheets do not. You can safely draw conclusions based on the  $r$ -value without the probability level providing the  $r$  value is not too small. A reasonable rule of thumb would be to assume that an  $r$  value of .4 or greater (in either the positive or negative direction) is fairly reliable.

## **SUMMARY**

Making sense of your data is really about organizing them into a meaningful story. As any good journalist would, begin by protecting your sources (client confidentiality). Then identify major themes in your lengthier notes (code wordy data) and organize all your information into a format that will be easy to construct the story from (prepare & organize). Before finalizing your story, verify your facts (check your data) and finally, build the story (analyze the data).

## **8—USING YOUR FINDINGS**

Research results that are not written up and shared are worth only the paper they are (not) written on. Regardless of the purpose of your evaluation project, stopping at the point of having gathered and analyzed data is stopping just short of the finish line. Even if you do not need to share the results with anyone, the process of "writing up" the project forces you to think through the meaning of your numbers. By doing so, you minimize the likelihood of drawing inappropriate conclusions. In fact, writing up your findings can be a very exciting step in the program evaluation process as it forces you to think not only about what you learned, but what the implications might be.

Writing a research report need not be difficult. To the contrary, it may be one of the simplest writing tasks you do because it typically follows a predictable formula. That means it does not require you to use the literary techniques that are needed even if only writing a letter to your friend. In fact, the best research reports are those that simply report the facts (what you did and what you found) and clearly tell the reader what conclusions you drew from them. But more about that in a while. First, we want to talk a bit more about *why* writing up your findings is important.

### **WHY WRITE IT UP?**

There are two critical reasons that you should take the time to report your study. One is because you owe it to yourself and the other because you owe it to the community. Let us explain.

**You owe it to yourself.** After all the effort you put into designing and implementing this study, you owe it to yourself to make sure the results are perceived accurately and taken seriously by those for whom they are intended. If you merely *talk* about what you found, even if you do a formal presentation, your message will easily be forgotten (or worse, distorted) and the actions you hoped to trigger will likely not happen. Human memory, unfortunately, is neither accurate nor long-lasting (unless, of course, it is about something you want people to forget in which case memory sticks like super glue!) Writing up your findings also helps you remember them.

Writing up your report ensures that the message you deliver comes across accurately and is in a format that does not fade with time. Not only does the permanence of the record

benefit your audience, but you would be amazed at how many details slip from your own mind once the study is over—details that you may want to revisit as you design future studies, evaluate policy changes, or simply answer questions about what you did.

***You owe it to the community.*** The moment you undertake an evaluation, no matter how large or small, you enter the scientific community. This means you become part of that group of people who chip away at the great, big block of *things we don't know*. Each researcher who gathers a chunk of data about a topic adds to the common knowledge. However, she or he does so *only* if he or she *reports* it. If the results are not written up and distributed, they do not contribute to our knowledge base. Unreported, if the findings are positive (demonstrated something worked), others will not be able to benefit from them. If the findings are negative (demonstrated something did not work), others will waste their time trying the same thing you did. Whether your community consists of the handful of people who work in your facility or an international network of researchers, having a written summary of your work lets you share your knowledge more easily.

## **WHAT NOT TO DO**

Don't overestimate the power of your study. If you conducted a project to, for instance, evaluate the usefulness to your therapeutic staff of a new diagnostic instrument for depression, make sure your conclusions reflect that. When you write up (or talk about) the study, talk about what your staff reported about use of the instrument with depressed clients in your facility. Do not over-generalize and conclude the instrument will have the same effect for all clients in your facility nor necessarily on depressed clients in other facilities. Draw your conclusions based on the study you did and the data you got. Report, but don't exaggerate, what you found.

## **WHAT TO DO AND HOW TO DO IT**

As we said earlier, writing a research report may be one of the easier writing tasks because it has a set structure into which you simply plug your information. The following discussion will lay out the formal structure that is used when scientists publish their work. Most likely, you will not need to include each section exactly as we show it. However, becoming familiar with all the pieces will help you organize

your own reports in a clear and easy to read manner. (Understanding this format will also make it easier for you to browse the research literature should you wish to.)

As you proceed, keep in mind who your audience is and write accordingly. Remember, the first and foremost objective of your report is to clearly communicate your findings. Thus, use language that is comfortable to you and familiar to your audience. If your project did not include components that we talk about in the following segment, just skip those sections. We are giving you the full, formal report structure to be sure we cover all the bases you may need covered.

## **THE PARTS OF A REPORT**

### **Executive Summary or Abstract**

Depending upon where you work, you will be familiar with one title or the other. Typically, "Executive Summary" is used for work-place reports and "Abstract" is used in the scientific literature. Regardless of the label, they both mean "summary of the report highlights." Basically, this short summary gives the reader a handy overview of what research question you addressed, how you addressed it, what you found, and what you concluded your findings mean.

You can build your abstract by answering these 4 questions:

1. What question(s) did you address?
2. What did you do to answer the question(s)?
3. What did you find?
4. What do you conclude about your findings?

For example, Brown & Miller's study on motivational interviewing (described in Chapter 3) might be abstracted as follows. The numbers in parentheses show you where the abstract answers each of the 4 questions.

---

#### **Sample Abstract**

(1) Twenty-eight clients participated in a study to address the question of whether motivational interviewing prior to substance abuse treatment improves outcome. (2) All consecutive intakes to a residential substance abuse program were alternately assigned to one of two conditions. Participants in the Motivational Interview (MI) group received two interview sessions prior to beginning treatment and participants in the No Motivational Interview (No-MI) group went directly to treatment. (3) Based on both client and therapist reports, participants in the MI group participated in their program to a greater degree than did participants in the No-MI group and showed significantly lower drinking levels at the three month follow-up. (4) We concluded that motivational interviewing prior to beginning treatment had a positive effect on outcome and merits additional research.

---

## Introduction

This section is where you describe the specific question your project is designed to address and why you are interested in it. If you have any predictions (hypotheses) about what the answer will be, this is where you express them. When stating predictions, it is a good idea to explain *why* you believe they will occur. Not only does it help your reader understand your line of thinking, it also helps you figure out where and why you were surprised if the results do not turn out as you expect.

The length of the Introduction will vary with the project it is introducing. Depending upon the complexity of your project and the intended audience for your report, the Introduction can be as short as a few sentences or many pages in length.

## Methods

The Methods section is where you precisely describe how you went about answering your research question. It is, in effect, the *recipe* for your project. You can test how well your Methods section is written by letting someone unfamiliar with the study read it and then describe to you what you did. If s/he gives a pretty accurate description, your description is clear. If not, consider cleaning it up a bit.

It is easiest to put your Methods section together by structuring it in subsections as shown in the following table. Again, if this is more detailed than your project warrants, only use the pieces that are relevant.

<b>Sections Within the Methods Section</b>	
Participants	Describe your participants. How many males and females participated in the study?  If there are client characteristics that are relevant to the study, describe them here. (For example: average age, education level, presenting problem, chronicity of problem, and so on.)
Materials	If you used questionnaires, tests, surveys, or other materials to gather data, this is the place to describe them. Typically, you would provide the name of the instrument, its purpose, the number of questions it contains, the format of the questions (e.g., multiple choice), and what the range of possible scores means.

<b>Sections Within the Methods Section</b>	
	<p>For example, you might describe the <i>Short Inventory of Problems</i> like this.</p> <p><i>Short Inventory of Problems (SIP-2L). This 15 item yes/no questionnaire asks respondents whether various problems associated with alcohol use have ever happened to them. Total SIP-2a scores range from 0, indicating no problems, to 15, indicating a high level of associated problems.</i></p>
Procedure	<p>If you think of the Material as the ingredients for your cake, the Procedure is the instructional part of the recipe. In this section you describe, in order, what you did. It is not necessary to construct an elegant literary piece that walks your reader through the project. Rather, simply state each step of the project. If it helps you to number the steps, do so.</p>

## Results

The Results section is the place you *report* your findings. You do not interpret them here nor ponder their meaning. You simply report them.

The easiest way to structure this section is to present as many of the data as possible in table form and use prose to summarize the findings. For example, consider the following results section for a hypothetical study that compared the rate of new HIV cases among intravenous drug users between two communities: one that offered a needle-exchange program and the other that did not. The number of new HIV cases in each community was tracked for one year. At the end of the year, the mean number of monthly cases was calculated and a t-test used to determine whether the means differed significantly.

---

### Sample Results

---

Table 1 summarizes the number of HIV positive reports in the experimental and control communities. As can be seen, the community with the Needle Exchange Program averaged almost 2 reports fewer per month ( $M = 1.03$ ) than did the control community ( $M = 3.0$ ).

Table 1. Summary of number of reports from the Needle Exchange Program and No Program communities.		
	Needle Exchange Program	No Program

Mean	1.03	3.00
Standard Deviation	1.21	1.37
Observed range	0 – 5	0 – 5

A t-test comparing the two means found them to be statistically significant ( $t(22) = -1.94, p < .05$ ).

---

## Discussion or Conclusion.

This is where you talk about what your findings *mean*. Did they support your predictions (hypotheses) or surprise you? What did you learn about your program? Do your findings suggest any changes that could be made? What might you do as a next step in your program evaluation? In essence, the Discussion (or Conclusion, they are used inter-changeably) is where you pull it all together and make recommendations for the future.

Using the sample needle exchange study described above, here's how a Discussion section might read. Bear in mind, each writer will naturally have his or her own slant on interpretation so you might well draw different conclusions from those we have drawn.

---

### Sample Discussion

This comparison suggests there may well be an important reduction in the number of HIV positive cases following introduction of a Needle Exchange Program. We observed a significantly lower number of reported HIV cases in the experimental community.

In spite of the positive results, there were a number of operational challenges that remain to be resolved before wide-spread implementation of this program is recommended. Paramount among these are getting word about the program out to the using community and consistently tracking clients who do avail themselves of this service.

It is our recommendation that the program be continued on an evaluative/developmental basis for the next year. The objective of this next phase of program evaluation will be to prepare the program for county-wide implementation at the end of that time.

---

## ***BUILD YOUR OWN REPORT***

Use Worksheets #12 through #16 beginning on page 140 to pull together the facts and your thoughts about the program evaluation study you did. If you are reading this chapter before you have all your data, just fill in the Introduction and Methods section for now. When your data are in and analyzed, you will be only two steps away from having a final report.

Notice that we have placed the Abstract worksheet last. This is, obviously, because it summarizes the information in the other sections of the report. So, although it appears first in your report, it is easiest to write it last since you can pretty much build it out of borrowed phrases and sentences from the main body of the report. Remember, though, this summary is a very important part of your report. It is the first section your reader sees and often determines whether or not s/he proceeds to read the balance of your report.

This page left intentionally blank.

WORKSHEET #12. PLAN YOUR REPORT — INTRODUCTION

**Introduction.** What is the topic of your study and what specific questions are you hoping to answer? Explain the reasoning behind your study. How do you think it will turn out?

**Methods.**

**Participants.** Describe the group of people who participated in your study.

**Materials.** Describe all the materials, such as surveys or questionnaires, used in the study.

**Procedure.** Describe in a step-by-step fashion what you did to conduct the study and gather the data.

WORKSHEET #14. PLAN YOUR REPORT – RESULTS

**Results.** Present the summary statistics that describe your data. Report any statistical analyses you performed and what their results were.

WORKSHEET #15. PLAN YOUR REPORT – DISCUSSION/CONCLUSION

**Discussion/Conclusion.** Did the findings support your predictions? What did you learn about your program? Do your findings suggest any changes that could be made? What might be a reasonable next step in your program evaluation?

WORKSHEET #16. PLAN YOUR REPORT – ABSTRACT

***Abstract.***

1. What did you want to learn?
2. What did you do to find out?
3. What did you find?
4. What did you conclude about your findings?

This page intentionally left blank.

## ***WHO TO SHARE IT WITH***

In deciding who should receive copies of your report, there are a few "musts" we can list (see Worksheet #17, page 148) but beyond that you are the best judge. Think about who in your facility could benefit from the information or might simply be interested (and perhaps could facilitate future projects). Also consider other facilities that might be thinking about making similar changes or adopting your protocol.

This page intentionally left blank.

WORKSHEET #17. DECIDE WITH WHOM TO SHARE YOUR REPORT

*The following people should be included in the list of those who receive copies of your report.*

Your boss

Other management who were involved in any way with the project

The funding agency for the project (if any)

Staff who participated in the project

Staff who did not participate but were inconvenienced (attach a thank you note)

Clients who participated and requested a copy of the final report

*List others in your organization or in other organizations who you think would value receiving a copy.*

This page intentionally left blank.

## **DIFFERENT REPORT FORMATS**

The report format we demonstrated in this chapter is the standard used most often to report research results. You will almost never go wrong if you rely on it. However, there may be times that your intended audience requires a particular format unlike the one presented here. By all means, present your results in the format that will be best accepted by your audience; but, do not do so at the expense of presenting the information inaccurately.

All of the bits and pieces that you made notes on in Worksheet #12 should remain part of your report. Regardless of how casual, or just plain different, the required format is, it must be complete. Your readers must still understand *what* your study was testing (the Introduction), *how* you tested it (the Methods), *what* you *found* (the Results), and *what* it *means* (the Discussion). Without all these pieces, the reader has no way of making sense of your project. You can, of course rearrange the pieces to meet your audience requirements, make the language as casual or formal as needed, and be as succinct or long-winded as you like. The bottom line, however, is be complete.

The following “mini-reports” illustrate just how misleading or useless a report can be without all the essential information.

<b>REPORT THAT IS MISSING ESSENTIAL “WHAT” INFORMATION (INTRODUCTION)</b>
The 50 item survey asked for opinions on the program’s effectiveness. The possible survey score ranged from 0 to 250 with a higher score indicating a more positive opinion. One hundred (100) community members were chosen at random as they shopped in the local mall and asked to complete the survey. Completed survey scores ranged from 50 to 200 with a mean score of 189 (SD = 23.5). The results indicate that the community is generally satisfied with the program.
<b>PROBLEM:</b> This report does not tell us why the survey was undertaken. Is this a new program that we want community reaction to or an old one that we are concerned has slipped in the community’s esteem? Is a mean score of 189 good or not? Have prior surveys of this sort or in this community or with this program obtained stronger or weaker scores? In a nutshell, we don’t know enough about what these findings mean to be pleased, displeased, or make any decisions.

**REPORT THAT IS MISSING ESSENTIAL “HOW” INFORMATION (METHODS)**

There has been some concern about how comfortable clients are with the new sign-in procedure at the clinic. To determine whether we need to revise this procedure, clients were asked as they entered the clinic whether they “liked,” “did not like,” or had “no opinion” about the procedure. All clients surveyed indicated they “did not like” the procedure. Based on this, we recommend that a task force be formed to redesign our sign-in procedure.

**PROBLEM:** Who surveyed the clients and how were they asked? How many clients were surveyed? If only 2 of the 100 clients who attended the clinic this week were surveyed or 80 of the 100 were surveyed, the results would have very different meanings. Assigning a task force (personnel costs), redesigning paper work (material costs), and possibly hiring new front-office staff (more personnel costs) should hardly be done on the basis of 2 people’s responses. On the other hand, if 80% of your clientele is unhappy, it is indeed important to address their concerns.

**REPORT THAT IS MISSING ESSENTIAL “FOUND” INFORMATION (RESULTS)**

Inclusion of a 2-session motivational interview prior to alcohol treatment was compared to treatment without the interview to determine whether the interview would improve clients’ participation in the program activities (i.e., attendance at group sessions, completion of homework assignments, attendance at individual sessions). New clients during the months of June and July were alternately assigned to either an Interview or No Interview condition and their program participation monitored until they either completed treatment or dropped out. We found that clients in the Interview condition participated in program activities to a greater extent than did No Interview clients and recommend the Interview be made a standard component of our alcohol program.

**PROBLEM:** Although you and your mother know you are a very careful thinker and would not draw conclusions based on weak data, the rest of the world is less trusting. Without a presentation of the actual results (How many group and individual sessions were attended? How many homework assignments were completed?), we do not know whether these conclusions are based on a weak, statistically insignificant lead by the experimental group or a strong, undeniable superiority.

**REPORT THAT IS MISSING ESSENTIAL “MEANS” INFORMATION (DISCUSSION)**

To evaluate the effectiveness of the new public health anti-smoking campaign, the new material replaced the current program in the town of Malhomme and the current program was left in place in the town of Bonhomme. Both campaigns ran for the first 10 months of 1995 and the following outcomes were monitored: # of people in each community who reported *starting* to smoke during that period, # of people who reported *attempting* to quit smoking, and # who reporting *successfully* quitting. At the end of the test period, we found that 20% fewer people in the test community (Malhomme) started smoking but 20% more people in the control community (Bonhomme) successfully quit. Attempted quit rates in the two communities were identical.

**PROBLEM:** What do these results mean? The reader can, we suppose, draw his or her own conclusions about whether it is more important to prevent new smokers or get old ones to quit. However, it is the researchers who are most intimate with the nuances of these programs and their results and thus, most competent to draw conclusions about what these results mean.

## **SUMMARY**

Results not reported are worth only the paper they are (not) written on. Whether you have conducted a large-scale, multi-site study or simply surveyed the clients who make up your work day, the information you gather is important and should be treated as such. You may not need to prepare a lengthy, formal report. Merely gather on paper the essential information that your study was built on (what, how) and generated (found, means). Make your report clear and to the point, and draw your conclusions based on the data you got.

If you feel tempted to stop your project short of the report step, consider how little we would know of our own history if archaeologists felt the same way. Think of all the thousands of individuals who have dug for fossils, glued together pottery shards, and excavated lost cities. What if they had not reported their individual “little” findings? If they had done their own thing, assembled their little pot or framed their ancient parchment for personal enjoyment, the wonderful tapestry of human cultural evolution would remain lost to us. The grandeur that was ancient Greece would remain hidden in a disjointed collection of private pottery collections. The splendor of King Tut’s tomb would be just one more gaudy display of wealth rather than a valuable piece in the puzzle of the cultural roots of Western Civilization. How much would be lost!

The page intentionally left blank.

## 9—WRAP-UP

Congratulations on having come this far. Whether you have already kicked off a program evaluation project in your facility or are simply completing your once-over on this manual, you have taken a significant step forward. You have begun to think about your program as a dynamic, evolving entity. Furthermore, you know how the nature of its evolution can be molded by careful program evaluation.

The central message we tried to convey is that program evaluation is not only critical to the ongoing success of a program, but is also simple enough to be undertaken by almost any agency. The secret to effective program evaluation is to carefully formulate your evaluation question and then methodically plan how you will answer it. The easiest way for you to do this planning is to follow the path laid out by the chapters and worksheets in this manual. If you have not already filled in the worksheets, the following summary will help you do so. If you have completed the worksheets, this summary will provide one final check for you to make sure you did not overlook any important steps.

Begin by thinking about what sort of information your program can most benefit from at this time. Do you need information about how much service you provide (*volume*) or how well that amount meets the community's needs (*reach*)? If you want to know or show how well your program works, evaluation of your program's *effect* would be in order. On the other hand, you might need to figure out how you can provide more or better service with the same budget. In this case, *value* research is the way to solve your problems.

WORKSHEET #1

Having identified the information you believe would be helpful, look carefully at what information you currently have. You may already have some of the answers you need. Those that you don't are good starting points for discussing program evaluation research projects with your colleagues.

WORKSHEET #2  
WORKSHEET #3

Narrow down the possible evaluation projects to one useful question that you and your colleagues can all get behind. However, do not make the mistake of assuming that just because everyone thinks it's a good idea,

WORKSHEET #4  
WORKSHEET #5  
WORKSHEET #6

they will support it all the way to its conclusion. Take the time to insure your project against abandonment and sabotage. Identify all the possible benefits as well as the potential barriers. Forewarned, as they say, is forearmed.

Armed, so to speak, with a meaningful evaluation question, a team, and a clear sense of where the highs and lows of this project are likely to occur, it is time to translate your general question into a *research question*. As you do, keep in mind that good research questions are very specific about what they ask. Your research question will guide the design of your study. It will help you figure out what your *dependent variables* will be and whether you need *independent* or *predictor variables*, or *control* and *experimental groups*. As you nail down the details of your design, give careful thought as well to the selection of your study *participants*.

WORKSHEET #7  
WORKSHEET #8

Selecting the instruments to measure your dependent variables must be done carefully. There are many, many available in the substance abuse field but not all are *reliable*, *valid*, and *interpretable*. If you are not well-versed in the measurement field, begin your search for good instruments with our recommendations and consult with your more knowledgeable colleagues or use the recommended readings in Chapter 4. As you choose your instruments, keep in mind what you want them to do for you. If your project is concerned with volume or reach questions, you may be particularly interested in the information available from *screening* or *diagnostic* instruments. For example, what percentage of people present with various diagnoses and what percentage of our staff is capable of treating those diagnoses? On the other hand, projects that look at effect or value may have more use for *diagnostic* and *assessment* instruments. Examine the instruments you select and make sure they measure the "thing" you are interested in and that they were appropriately developed for the population in which you will measure it. If you are conducting follow-up evaluation, the instruments should also address the appropriate timeframe as well as conditions that are indeed likely to change. For example, do not use an instrument that asks about lifetime drinking history if you are examining change in drinking during the 3 months since treatment ended.

WORKSHEET #9  
WORKSHEET #10

Gathering program evaluation data is the first half of getting answers to your questions. The second half is making sense of the information you gather and organizing it into a readable report.

The first step in this process is setting up your data-gathering process to ensure client confidentiality. Examine the data you will be

WORKSHEET #11

gathering, the questions you are asking, and decide what bits of information (numbers, usually) you will analyze and how. As you gather data, keep them organized so that when it comes time to analyze them, you will be ready to go.

Finally, summarize the work you have done and what you learned into a report that you can share with colleagues. The elements of a useful report are quite simple. Tell your audience what you wanted to find out (the *Introduction*) and how you set about doing that (the *Methods*).

WORKSHEET #12  
WORKSHEET #13  
WORKSHEET #14  
WORKSHEET #15  
WORKSHEET #16

Explain what you found (the *Results*) and what you concluded those findings mean (the *Discussion*). Top off this report with a short *Executive Summary* that will give your audience a quick overview of these four sections.

Consider with whom you will share your report. Obviously, your boss, team members, and funding sources are interested, but there may be others who could benefit from your findings. Are there other agencies contemplating changes similar to those you evaluated? Perhaps there are people in your own organization who were not

WORKSHEET #17

involved in the project but could learn from its outcome. Remember that sharing of knowledge is the surest way to move a field forward in a positive direction. It benefits yourself, your program, your clients, and ultimately, through improved services, the communities in which we live.



**ADDITIONAL READING**

If you are interested in furthering your understanding and skills in program evaluation, we recommend the following.

Addiction Research Foundation (1993). *Directory Of Client Outcome Measures For Addictions Treatment Programs*. Toronto: Addiction Research Foundation.

Lettierie D.J. (Ed.) (1987). *Research Strategies In Alcoholism Treatment Assessment*. New York: Haworth Press.

Lettierie, D.J., Allen, J., & Caldwell, F. (Eds.) (1992). *A Primer Of Research Strategies In Alcoholism Treatment Assessment*. NIAAA Treatment Handbook Series, No. 3. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

National Institute on Drug Abuse (1993). *How Good Is Your Drug Abuse Treatment Program? A Guide To Evaluation*. Rockville, MD: NIDA.

Shadish, W.R., Jr., Cook, T.D., & Leviton, L.C. (1991). *Foundations Of Program Evaluation: Theories Of Practice*. Newbury Park, CA: Sage Publications.

## **APPENDIX A—MEASUREMENT INSTRUMENTS**

This appendix is divided into two parts. In the first, you will find a convenient table of the ordering information for the copyrighted instruments discussed in Chapter 4 along with some of the public domain instruments. In the second, we have provided copies of the remaining public domain instruments discussed in that chapter.

### **ORDERING INFORMATION**

Note: Accurate as of June/95. Cost and ordering information may be subject to change.

Instrument	Approx. Cost	Ordering Information	Page Discussed
Alco-Sensor	\$600.	Intoximeters Inc., 1901 Locust Street, St. Louis, MO 63101. 1-800-451-8639. FAX: 314-241-9734.	68
Drug Abuse Screening Test (DAST)	\$5.50 Can. per 100 copies	Addiction Research Foundation 33 Russell Street Toronto, Ontario Canada M5S 2S1 1-800-661-1111	75
Structured Clinical Interview for DSM-Patient Edition (SCID-P)	\$35. to \$75. per 10-pack depending on package selected	American Psychiatric Press Washington, D.C. 1-800-368-5777	75
Diagnostic Interview Schedule (DIS)	Free	Dr. Linda Cottler 1 Brookings Drive Department of Psychiatry Washington University St. Louis, MO 63130-4899	75
Computer-Diagnostic Interview Schedule (C-DIS)	\$899 for 1st copy; 598/ea. for 2nd & 3rd; 299/ea. for additional	C-DIS Group 1800 Bank Street Ottawa, Ontario Canada K1V 0W3 1-800-465-8466 FAX: 613-526-0238	75
Manuals for: DrInC-2L DrInC-2R InDUC-2L InDUC-2R	Free	National Clearinghouse for Alcohol & Drug Information NIAAA 1-800-729-6686	78 78 78 79
Alcohol Dependence Scale (ADS)	\$15 Can. per 25-pack Starter's Kit	Addiction Research Foundation 33 Russell Street Toronto, Ontario Canada M5S 2S1 1-800-661-1111	79
Alcohol Use Inventory (AUI)	\$74. per 50-pack	National Computer Systems, Inc., P.O. Box 1416	80

Instrument	Approx. Cost	Ordering Information	Page Discussed
	Starter's Kit	Minneapolis, MN 1-800-627-7271	
Drug Use Self-Report (DUSR)	Facility license or individual test purchase	Center for Addictions Research & Evaluation 5460 Ward Road Arvada, CO 80002 303-421-1261	80
Addiction Severity Index (ASI)	Free	1-800-238-2433	81

## **REPRINTS OF PUBLIC DOMAIN MATERIAL**

Many of the following assessment instruments are available by visiting our website:

<http://CASAA.UNM.EDU>

Instrument	Page Discussed
Alcohol Use Disorders Identification Test (AUDIT) Scoring Rules for the AUDIT Screening Questionnaire	74
Quantity/Frequency/Volume-30 (QFV-30)	77
Quantity/Frequency/Volume-90 (QFV-90)	77
Quantity/Frequency Interview	77
Drinkers Inventory of Consequences-2L (DrInC-2L) DrInC Scoring Sheet DrInC Profile Sheet for Men DrInC Profile Sheet for Women <i>Note: Ordering information for Manual above.</i>	78
Drinkers Inventory of Consequences-2R (DrInC-2R) DrInC Scoring Sheet DrInC Profile Sheet for Men DrInC Profile Sheet for Women <i>Note: Ordering information for Manual above.</i>	78
Inventory of Drug Use Consequences-2L (InDUC-2L) InDUC Scoring Sheet <i>Note: Ordering information for Manual above.</i>	78
Inventory of Drug Use Consequences-2R (InDUC-2R) InDUC Scoring Sheet <i>Note: Ordering information for Manual above.</i>	79
Goal Attainment Scale (GAS) GAS Follow-up GAS Instructions	79
Severity of Alcohol Dependence Questionnaire (SADQ) SADQ Scoring Instructions SADQ Scoring Sheet	80

## **APPENDIX B—ADOLESCENT & SPANISH INSTRUMENTS**

Here you will find a convenient table of ordering information for copyrighted instruments appropriate for use with adolescents as well as copies of public domain instruments for both adolescents and Spanish-speakers.

### **ORDERING INFORMATION**

Note: Accurate as of June/95. Cost and ordering information may be subject to change.

Instrument/purpose	Cost	Ordering Information	Page Discussed
Adolescent Drinking Index (ADI)	\$48/Kit: (manual +25-pack	Psychological Assessment Resources Odessa, FL 1-800-331-TEST	82
Drug Use Screening Inventory (DUSI)	Free	Dr. R.E. Tarter Dept. of Psychiatry University of Pittsburgh Medical School 3811 O'Hara St. Pittsburgh, PA 15213	82
Rutgers Alcohol Problem Inventory (RAPI)	Free	Dr. H.R. White Center of Alcohol Studies Rutgers University Piscataway, NJ 08855	83
Adolescent Drug Abuse Diagnosis (ADAD)	Free	Dr. A. S. Friedman Belmont Center for Comprehensive Res. 4200 Monument Rd. Philadelphia, PA 19131 215-877-2000	83

### **REPRINTS OF PUBLIC DOMAIN MATERIAL**

Website address: <http://CASAA.UNM.EDU> \*

Instrument	Page Discussed
Adolescent Alcohol Involvement Scale (AAIS) AAIS Scoring Instructions ( <i>see following pages</i> )	82
* Inventario Sobre las Consecuencias del Use de Alcohol-2L (DrInC-2L)	88
* Inventario Sobre las Consecuencias del Use de Alcohol-2R (DrInC-2R)	88

**ADOLESCENT ALCOHOL INVOLVEMENT SCALE  
(AAIS)**

1. How often do you drink?

- a. never
- b. once or twice a year
- c. once or twice a month
- d. every weekend
- e. several times a week
- f. every day

2. When did you have your last drink?

- a. never drank
- b. not for over a year
- c. between 6 months and 1 year ago
- d. several weeks ago
- e. last week
- f. yesterday
- g. today

3. I usually start to drink because:

- a. I like the taste
- b. to be like my friends
- c. to feel like an adult
- d. I feel nervous, tense, full of worries, or problems
- e. I feel sad, lonely, sorry for myself

4. What do you drink?

- a. wine
- b. beer
- c. mixed drinks
- d. hard liquor
- e. a substitute for alcohol—paint thinner, sterno, cough medicine, mouth-wash, hair tonic, etc.

5. How do you get your drinks?

- a. supervised by parents or relatives
- b. from brothers or sisters
- c. from home without parents' knowledge
- d. from friends
- e. buy it with false identification

6. When did you take your first drink?

- a. never
- b. recently
- c. after age 15
- d. at ages 14 or 15
- e. between ages 10–13
- f. before age 10

7. What time of day do you usually drink?

- a. with meals
- b. at night
- c. afternoons
- d. mostly in the morning or when I first awake
- e. I often get up during my sleep and drink

8. Why did you take your first drink?

- a. curiosity
- b. parents or relatives offered
- c. friends encouraged me
- d. to feel more like an adult
- e. to get drunk or high

9. How much do you drink with you do drink?

- a. 1 drink
- b. 2 drinks
- c. 3–6 drinks
- d. 6 or more drinks
- e. until “high” or drunk

10. Whom do you drink with?

- a. parents or relatives only
- b. with brothers or sisters only
- c. with friends own age
- d. with older friends
- e. alone

11. What is the greatest effect you have had from alcohol?

- a. loose, easy feeling
- b. moderately “high”
- c. drunk
- d. became ill
- e. passed out
- f. was drinking heavily and the next day didn’t remember what happened

12. What is the greatest effect drinking has had on your life?

- a. none—no effect
- b. has interfered with talking to someone
- c. has prevented me from having a good time
- d. has interfered with my school work
- e. have lost friends because of drinking
- f. has gotten me into trouble at home
- g. was in a fight or destroyed property
- h. has resulted in an accident, an injury, arrest, or being punished at school for drinking

13. How do you feel about your drinking?

- a. no problem at all
- d. I often feel bad about my drinking

- b. I can control it and set limits on myself
- c. I can control myself, but my friends easily influence me
- e. I need help to control myself
- f. I have had professional help to control my drinking

14. How do others see you?

- a. can't say, or a normal drinker for my age
- b. when I drink I tend to neglect my family or friends
- c. my family or friends advise me to control or cut down on my drinking
- d. my family or friends tell me to get help for my drinking
- e. my family or friends have already gone for help for my drinking

**ADOLESCENT ALCOHOL INVOLVEMENT SCALE  
(AAIS)**

**SCORING INSTRUCTIONS**

The highest total score is 79. An *a* response is scored 1 (except on questions 1, 2, 6, 12, 13, and 14, on which *a* = 0); *b* = 2; *c* = 3; and so on to *h* = 8. When more than one response is made, the one with the higher or highest score is used. An unanswered question is scored 0.

The AAIS conceptualizes adolescents' alcohol misuse as drinking to the extent that it interferes with anyone or any combination of three areas: psychological functioning, social relations and family living. A respondent's total score places him somewhere on a continuum from 0 (abstainer) to 42–79 (misuser).

From: J. Mayer & J. Filstead (1979) The Adolescent Alcohol Involvement Scale: An instrument for measuring adolescents' use and misuse of alcohol. *J. of Studies on Alcohol*, 40 (3), 291–300.

## ***APPENDIX C—MANUAL EVALUATIONS***

Thank you for taking the time to evaluate this manual. Your feedback will be helpful to us for improving this manual.

Please complete and return Evaluation Form #1 when you have finished reading this manual for the first time. Complete and return Evaluation Form #2 after you have used the information from this manual to initiate a program evaluation research project in your facility.

This page intentionally left blank.

## **EVALUATION FORM #1**

Please return this form by mailing it to Dr. William Miller, Dept. of Psychology, Logan Hall, University of New Mexico, Albuquerque, NM 87131-1161.

Answer each of the following questions by circling the number of the one statement that best describes your opinion.

### **How easy or difficult to read was the manual?**

1. *Very easy* to read
2. *Somewhat easy* to read
3. *Neither* easy nor difficult
4. *Somewhat difficult* to read
5. *Very difficult* to read

### **How much of the information in this manual was useful?**

1. *All* of the information was useful.
2. *Most* of the information was useful.
3. *Some* of the information was useful.
4. *Very little* of the information was useful.
5. *None* of the information was useful.

### **How easy or difficult to understand was the information in this manual?**

1. *Most* of the information was *extremely easy* understand.
2. *Most* of the information was *somewhat easy* to understand.
3. *Some* of the information was *easy* to understand.
4. *Most* of the information was *somewhat difficult* to understand.
5. *Most* of the information was *extremely difficult* to understand.

Please continue on the next page.

**Evaluation Form #1—continued**

---

Here is a list of all the chapters in the book. If any were particularly helpful or particularly difficult to read or understand, put a check in the column next to them. If you did not find any of them more helpful or difficult than the rest, skip this question.

	Particularly Helpful	Particularly Difficult
1—What is Program Evaluation Research?		
2—Benefits, Barriers, and Strategies.		
3—Questions and Methods		
4—Measurement Tool		
5—Assessing Special Populations		
6—Follow-up Evaluation		
7—Making Sense of Your Data		
8—Using Your Findings		
9—Wrap-Up		

How confident are you that you and/or your program facility will begin a program evaluation research project within the next 6 months?

1. *Very we **will** begin a project*
2. *Somewhat confident we **will** begin a project*
3. *Have no idea whether we will or will not begin a project*
4. *Somewhat confident we will **not** begin a project*
5. *Very we will **not** begin a project*

Please tell us what you liked best about this manual.

Please tell us what you liked least about this manual.

How can the manual be improved?

Thank you for your help.

## EVALUATION FORM #2

After you have begun (or completed) a research evaluation project, please complete and return this survey to Dr. William Miller, Dept. of Psychology, Logan Hall, University of New Mexico, Albuquerque, NM 87131-1161.

Answer each question by circling the number beside the statement that best describes your experience and opinion.

### How well did reading the manual prepare you for the realities of doing program evaluation?

1. I was *very well-prepared*.
2. I was *somewhat well-prepared*.
3. I was *not well-prepared*.

### How easy or difficult was it to apply the ideas in this manual to “real life”?

1. It was *very easy* to apply them.
2. It was *somewhat easy* to apply them.
3. It was *neither easy nor difficult* to apply them.
4. It was *somewhat difficult* to apply them.
5. It was *very difficult* to apply them.

### How easy or difficult was it to locate information in this manual?

1. It was *very easy* to find what I was looking for.
2. It was *somewhat easy* to find what I was looking for.
3. It was *neither easy nor difficult* to find what I was looking for.
4. It was *somewhat difficult* to find what I was looking for.
5. It was *very difficult* to find what I was looking for.

### How easy or difficult was it to actually get the evaluation project off the ground in your facility?

1. Very easy
2. Somewhat easy
3. Neither easy nor difficult
4. Somewhat difficult
5. Very difficult

Please continue on the next page.

### Evaluation Form #2—Continued

---

#### Which statement best describes the type of evaluation project you did or are doing? Circle all that apply.

1. *Program volume*—The project focus is/was on measuring how much traffic or activity occurs in our program.
2. *Program reach*— The project focus is/was on measuring how well our program resources and activities meet the needs of our client community.
3. *Program effect*—The project focus is/was on measuring how well the program accomplishes its intended purpose.
4. *Program value*—The project focus is/was on measuring how much effect you get for each program dollar.
5. *Other*, please describe \_\_\_\_\_

#### What dependent variable(s) does/did your project measure?

**Approximately how many people work in your agency? \_\_\_\_\_**

**How many staff are/were involved in the project? \_\_\_\_\_**

**What is the current status of your project?**

1. Just begun
2. Past the beginning but less than half complete
3. About half complete
4. More than half complete
5. Complete

**Given your experience so far, how likely do you think it is your agency will do future program evaluation projects?**

1. *Very likely* we will do additional projects
2. *Somewhat likely* we will do additional projects
3. *I really can't say* at this time.
4. *Somewhat unlikely* we will do additional projects
5. *Very unlikely* we will do additional projects

Please continue on the next page.

**Evaluation Form #2—Continued**

---

**What changes to this manual would have improved its usefulness to you as you worked on your project?**

**If your project is complete, please describe its outcome or enclose a copy of your report.**

Thank you for your help.