

Program Evaluation:

The Why, When, and How of Outcome Measures

Dr. Sarah Shelton, Psy.D.
Hope House for Women, Inc.
Augusta, Georgia



Introduction to Program Evaluation

- **Program Evaluation** is a formalized approach to studying the goals, processes, and impacts of projects, policies and programs. It reveals the intended and unintended effects of a program both positive and negative.

- **Outcome Measures** are tools used in program evaluation that show the effects of the program being examined.
 - Successes
 - Short-Comings
 - Guide for Future Direction

- **Philosophy of Program Evaluation is**
 “No information is bad information!”



WHY do a Program Evaluation?

- To determine the effectiveness of a program.
- To have objective data to show that a program is successful.
- To identify specific strengths and weaknesses of a program.
- To improve or modify a program.
- To increase competitiveness for grants. \$\$\$
- To contribute to the field of knowledge about specific programs or interventions = research.
- To guide staff training.



Why:

The History of Program Evaluation

- Program Evaluation became popular after the 1960's. The 60's were an era of initiating and mobilizing social programs. A great deal of money was spent on societal programs but there was no data on their effectiveness.
- Cost-Benefit Analysis: Are the resources needed to implement and maintain a program justified based on the program's results?



WHEN should you do Program Evaluation?

- To run a pilot program to test impact and to troubleshoot before implementing.
- To compare competing models of programs before selecting one for implementation.
- To evaluate a new program for effectiveness.
- To evaluate an old program to determine if it should remain in tact, be altered, or discarded.
- When you *know* it's good but need proof of efficacy.



HOW

To Do a Program Evaluation Project Overview...

- Internal vs. External Evaluator
- Relationships and Interpersonal Dynamics
- Choosing your Variables
- Selecting Outcome Measures
- Collecting Data
- Statistical Analysis
- Interpreting Results
- Communicating Findings
- What Next?



HOW:

Internal vs. External

Internal - Agency or representative conducts the program evaluation


pros: inside knowledge and understanding of program dynamics can prove helpful, established rapport with staff and patients, less expensive than hiring a consultant

cons: personal investment in outcome can lower objectivity, lack of a “fresh eye,” **most organizations lack internal professionals appropriately trained to design, analyze and interpret outcomes data.**



HOW:

Internal vs. External

 **External** – independent consultant brought in to complete specific task of program evaluation

pros: increased objectivity, new perspective, increased competency and access to program evaluation resources that may be out of reach for an organization (e.g. expensive statistical software).

cons: lacks “insider” knowledge and understanding of organizational dynamics, expensive \$\$\$, can be viewed as an “intruder” by staff and patients.



HOW:

Relationships & Interpersonal Dynamics

- Understanding the role of the program evaluator is key: Staff Training & Education.
- Clearly define who the “client” is upfront.
- Accept initial anxiety, resistance, and defensiveness as natural.
- Take the initiative to get to know the “players” in the game – conduct focus groups, ask for meetings, become a familiar face.
- Become an “Integrated Observer.”



HOW:

Choosing Your Variables

- What is important to the identified client?
 - Ex. – # of those completing program vs. those that drop out
- What is important to you?
- Identify tangible outcomes first
 - Ex. – Depression...BDI
- KEEP IT SIMPLE!






HOW:

Selecting Outcome Measures

Standardized vs. Created Measures

Standardized

-  Normed and validated with research to show reliability and validity
-  Readily available but various requirements for licensure exist depending on the measure (competency required, sometimes additional training necessary)
-  Can be costly



HOW:

Selecting Outcome Measures

Created

- Greater control over what you are measuring – *or so you think!*
- Less expensive than standardized measures.
- No restrictions on who can administer and score.
- No data on reliability, validity.



HOW:

Collecting Data

- Who is the collector?
- Pragmatic Issues – Scheduling, Organizational Priorities, Emergencies
- Confidentiality vs. Privacy
- Handling Resistance and Noncompliance
- Single vs. Group Administration
- Data Storage – HIPPA!
- Paper-pencil vs. Computer Entry



HOW: Statistical Analysis

■ Simple vs. Complex

■ Statistical Analysis Programs

■ Excel, Ecstatic, **SPSS**

■ Skill level of Evaluator

■ Intended Audience

■ Types of Variables

■ Descriptive, Nominal, Dichotomas, Categorical, Continuous

■ Individual vs. Aggregate Data



HOW: Interpreting Results

Be careful!

Were there enough participants? (Power, n)

Ex. Comparing pre and post program

Pre-Program n=50, Post-Program n=6

Was the range adequate? (score 1-10 vs. 1-3)

Ex. Outcomes show no significant change

Continuous score 0-100 or

Categorized as low (1), medium (2), high (3)

Are there any confounding or extraneous variables?

Ex. Outcomes of new program correspond with complete turnover of clinical staff

Research Integrity



HOW: Communicating Results

- Be honest but constructive in feedback
“Don’t OPPOSE until you can PROPOSE”
Be Solution Focused!
- Appreciate client’s anxiety and defensiveness about results
- Acknowledge the strengths and weaknesses of your evaluation
- Use language that audience understands
 - ❖ “A post-hoc t-test on the ANOVA revealed a marginal but statistically significant difference between the two groups.” NO!
 - ❖ “Females who participated in a yoga class were less depressed than females who did not participate in the yoga class upon discharge.” YES!



WHAT NEXT?

💰 \$

📄 Grants

Use program evaluation data to compete for grant funding for your program/organization.

Grants are available to fund the program evaluation itself.

Evidence-Based Practice



WHAT NEXT?


Continued

 3 P's

 PRESENT!

 PUBLISH!

 PROMOTE!

 Scientific and Professional Community are very interested in Program Evaluation as well as Evidence Based Practice – hot topic!



Sample Program Evaluation Project

 ***Hope House for Women, Augusta, Georgia***

Outcomes Study over 24-month Period

Population: Residential population of women with substance abuse /dual diagnosis – emphasis on homeless women who are pregnant or who have children.

Program: Multi-disciplinary long-term residential substance abuse treatment for women and their children.



Sample

Program Evaluation Project Continued

Variables of Interest - Outcome Measures

Psychological Distress – BSI GSI

PTSD – Modified PTSD Symptoms Scale

Substance Abuse – SASSI-3

Recognition of Problem – SOCRATES

Motivation for Change – SOCRATES



Sample

Program Evaluation Project Continued

Data Collection

Data were collected by group administration at 5 time points: admission, 1 month, 3 months, 6 months, and 12 months in the program.

Data Analysis

A series of repeated measures analysis of variance (ANOVA) were performed to assess for significant changes in residents' symptoms/distress, risk-behavior, and motivation for change across five measurement points. Analyses were performed using SPSS statistical software with an *a priori* alpha of $p < .05$.

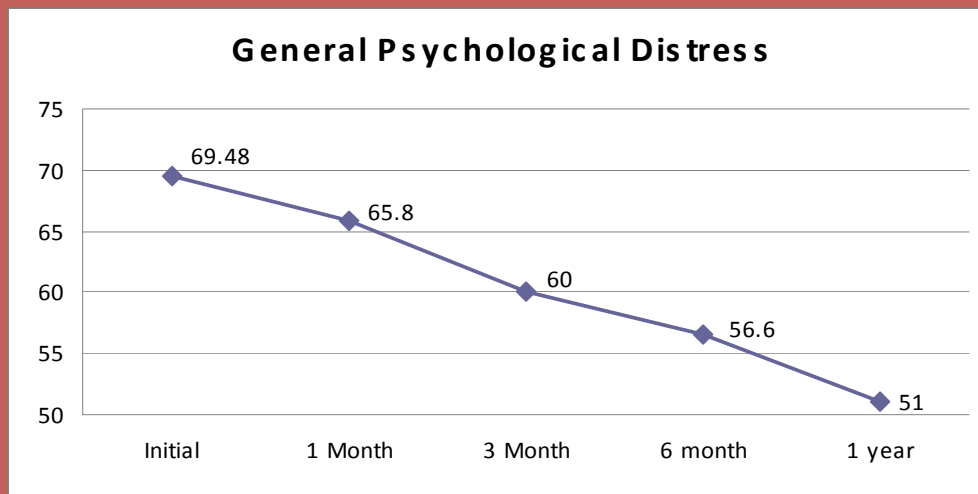


Sample

Program Evaluation Project RESULTS

General Psychological Distress - Brief Symptom Inventory Global Symptom Index (BSI)

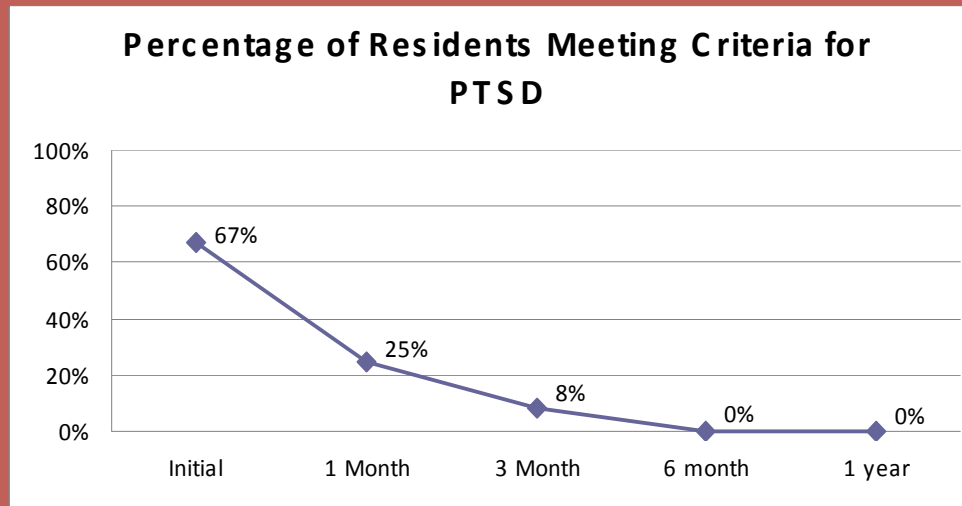
A significant reduction in scores on the BSI from baseline to the 3-month follow-up was observed ($N = 13$; $F = 69.34$, $p < .001$). This drop remained significant at both the 6-month ($N = 5$; $F = 27.44$, $p < .001$) and 1-year follow-ups ($N = 3$; $F = 192.50$, $p < .001$). The mean BSI score at the 3-month, 6-month, and 1-year follow-ups were at or below the standard cutoff for clinical significance.



Sample

Program Evaluation Project RESULTS

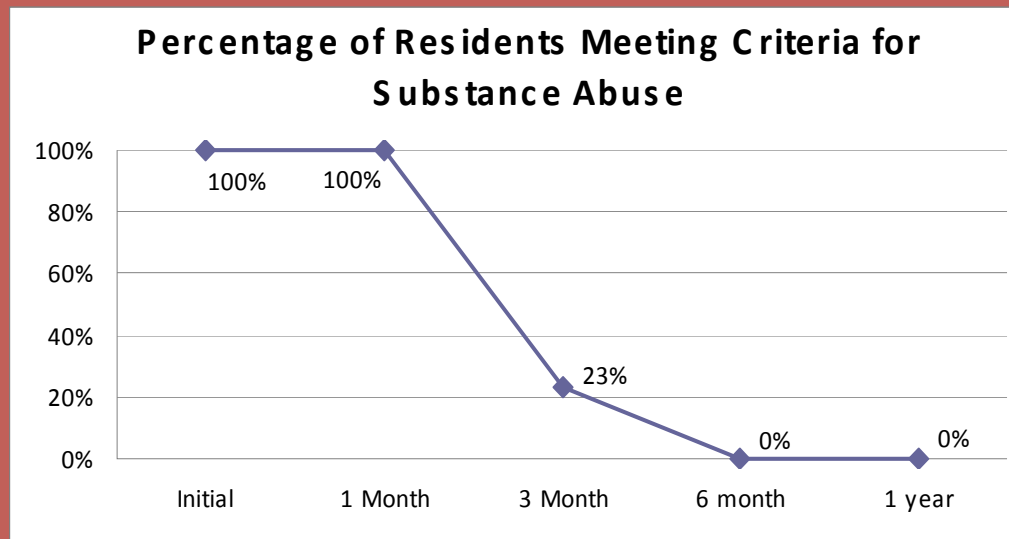
- **DSM-IV Criteria for PTSD** - Modified PTSD Symptom Scale (MPSS-SR)
- A significant reduction in the percentage of residents meeting DSM-IV criteria for PTSD was observed from baseline to the 3-month follow-up ($N = 13$; $F = 6.72$, $p < .01$), with only 8% of residents meeting criteria by the 3-month follow-up. This downward trend continued through the 6-month and 1-year follow-ups.



Sample

Program Evaluation Project RESULTS

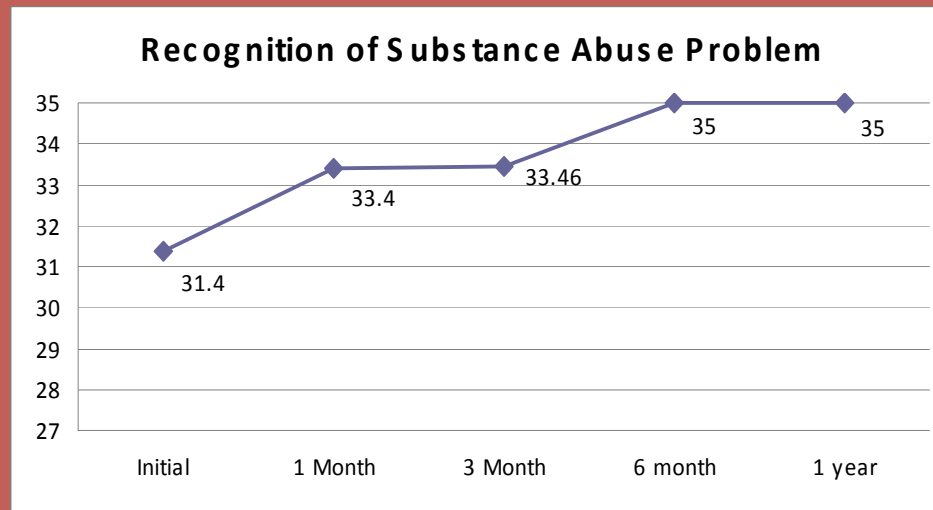
- **DSM-IV Criteria for Substance Abuse** - Substance Abuse Subtle Screening Inventory (SASSI)
- All residents met criteria for a substance abuse disorder at both the initial assessment and 1-month follow-up. However, a significant reduction in the percentage of residents meeting DSM-IV criteria for a substance abuse disorder was observed by the 3-month follow-up ($N = 13$; $F = 40.00$, $p < .001$), with only 23% of residents meeting criteria at Time 3.



Sample

Program Evaluation Project RESULTS

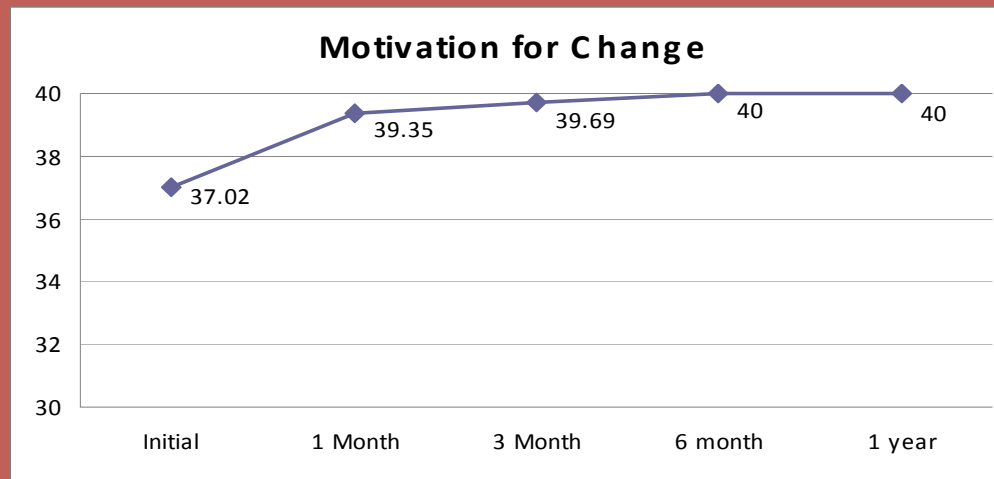
- **Recognition of Substance Abuse/Dependency** - Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES)
- No significant change in residents' acknowledgement of a substance abuse/dependence problem from the initial to the 3-month, 6-month, or 1-year assessments was observed. A slow (non-statistical) trend toward an increase in acknowledgement of a substance abuse/dependence problem is observed, with the greatest increase (6%) occurring between the initial and 1-month assessments.



Sample

Program Evaluation Project RESULTS

- **Motivation for Change** - Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES)
- A significant improvement in residents' motivation to take behavioral steps toward change was observed from the initial to the 3-month follow-up ($N = 13$; $F = 8.40$, $p < .01$). Similar to the pattern observed in participants' recognition of change, the greatest increase was observed between the initial assessment and the 3-month follow-up (6.3%) Although this improving trend continued through the 6-month and 1-year follow-ups, the change did not remain statistically significant through the final two observations.



Sample Program Evaluation Project

STRENGTHS

Longitudinal – 24 month time span

Consistency between program goals and outcome variables

Consistency between results and anecdotal evidence/clinical impressions or resident progress

CHALLENGES

Pragmatic issues with data collection


Statistical challenges - significantly more women enter in program than remained at post-admission data collection points

Aggregate data shows overall trends but not clinically useful for individual patients

FUTURE DIRECTIONS: Expand program evaluation data to include 1.) information on patients at discharge 2.) outcomes for children of the residents 2.) clinically useful information for individual patients vs. aggregate data



CONCLUSION

 Program Evaluation is scientifically and professionally necessary to promote and provide evidence-based practice in the field of mental health care.

Clinical Advantages: ensures quality of care for mental health consumers.

Professional Advantages: promotes the field of mental health as a science, demands provider accountability, helps standardize treatments.

Social & Political Advantages: helps fight to end discrimination for substance abuse and mental health services through social-political efforts such as mental health parity in insurance coverage, allowing us to help more people in need.

