

EVALUATION

Evaluating Effectiveness of Health Education Programs

Presented by: Richard T (Tim) Patton, Director
Professional Master of Public Health & Institute
for Community Health and Minority Medicine,
School of Graduate Medical Science, BARRY University

Uses for Evaluation:

- ◆ **Program objectives met.**
- ◆ **Document strengths and weaknesses.**

Uses for Evaluation:

- ◆ **Monitor standards of performance and quality control.**
- ◆ **Provide data for fiscal accountability.**

Uses for Evaluation:

- ◆ **Improve staff member skills.**
- ◆ **Meet grant or contract requirements.**

Uses for Evaluation:

- ◆ **Promote public relations and awareness.**
- ◆ **Determine if program is generalizable.**

Uses for Evaluation:

- ◆ **Contribute to knowledge base of public health education program design.**
- ◆ **Identify hypothesis for future evaluation.**

Types of evaluation

Preferred term

- **Needs Assessment**
- **Process Evaluation**
- **Impact Evaluation**
- **Outcome Evaluation**

Other Term

- Diagnostic Evaluation**
- Formative Evaluation**
- Summative Evaluation**
- Summative Evaluation**

Evaluation

Levels of Evaluation

- ◆ **Process** = Evaluating the Program's process or activities
- ◆ **Impact** = Evaluating the Behavior Change
- ◆ **Outcome** = Evaluating the Health Status

Levels of Evaluation

- ◆ **Process** – something changes as a result of **planned** learning and management **activity**
- ◆ **Impact** – intervention leads to an **observable behavior** impacts on the health status
- ◆ **Outcomes** – behavioral adaptation leads to an improvement in **health status**

Evaluation Progression

Process

Program

Instructors
Content (learning objectives)
Methods
Time Allotment

Impact

Behavior

Knowledge → Behavior
Attitude change
Habit/Skill development

Outcome

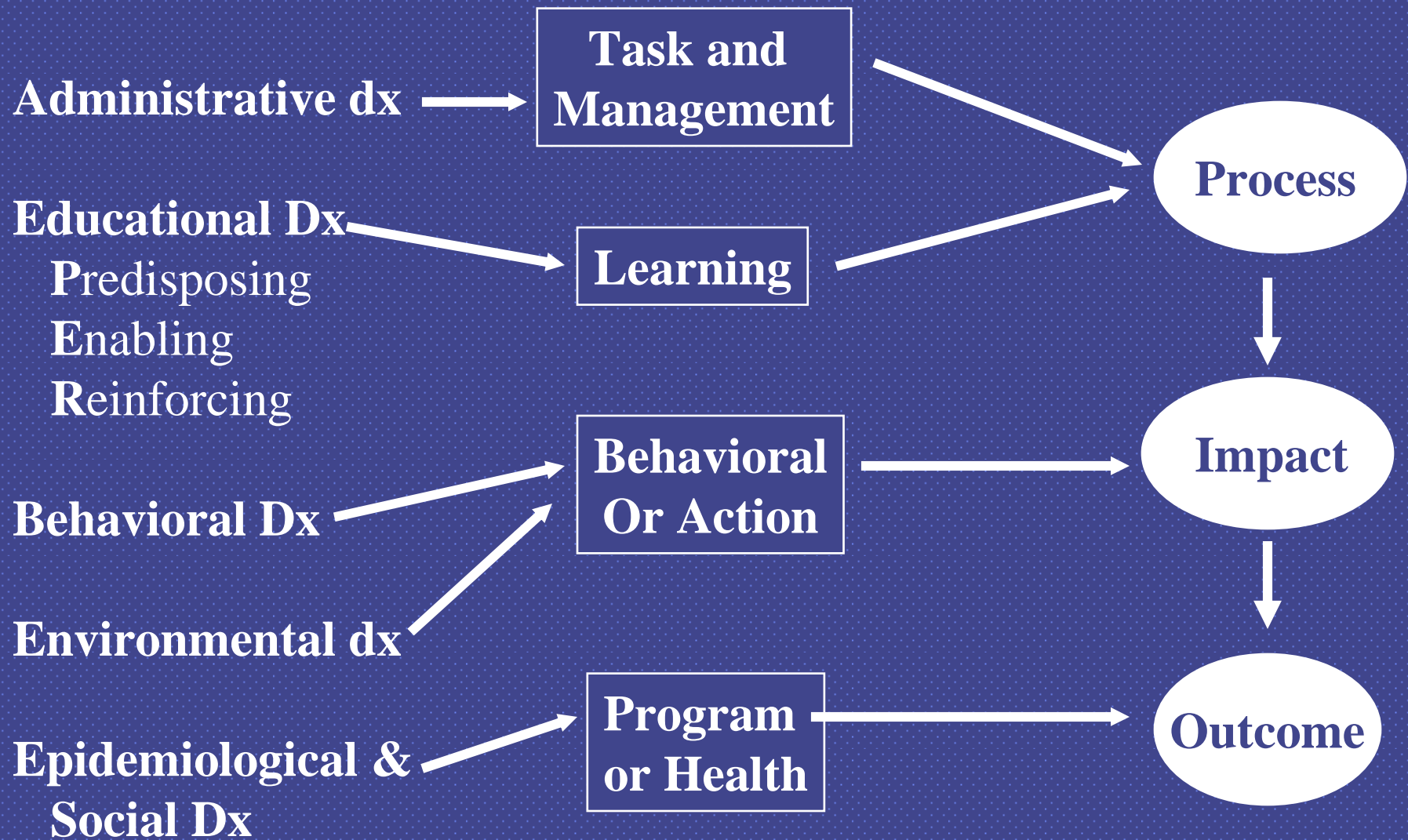
Health

Mortality
Morbidity
Disability
Quality of life

Precede Steps

Objective Levels

Evaluation Levels



Types of Evaluation

- ◆ **Formative** – ongoing improvement in an intervention program or curriculum
- ◆ **Summative** – assesses the extent to which a finished product or program causes change in the desired direction in the target population

Evaluation Designs

- ◆ Same type of designed used to conduct research are used to conduct evaluations
- ◆ Designs are intended to measure and determine if **change** is a result of the **intervention**.

Types of Design

- ◆ Non-experimental or single group design
- ◆ Quasi-experimental design
- ◆ Experimental design

Non-experimental (or single group design)

- ◆ Does *not* use experimental or control group
- ◆ Participants *not* randomly assigned
- ◆ Data can be collected from participants at end of program or at beginning and end, and compared for differences using pre and post test
- ◆ Cannot determine if changes are result of the intervention because of decreased control

Quasi-experimental design

- ◆ Experimental group and a control group or comparison group are formed by means other than random assignment
- ◆ Data is collected from both prior to and after the intervention
- ◆ Lack of random assignment decreases control
 - Hence, changes in the experimental group may support the effectiveness of the intervention

Experimental design

- ◆ Use of random assignment of participants into control and experimental groups
- ◆ Data collected from both groups before and after intervention
- ◆ Changes in the experimental group are the best evidence of the effectiveness of the intervention because of increased control

Evaluation Designs

- ◆ **Statement of objectives**
- ◆ **Definition of data to be collected**
- ◆ **Methodology**
- ◆ **Instrumentation**
- ◆ **Data collection**
- ◆ **Data processing**
- ◆ **Data analysis**
- ◆ **Reporting**

Statement of Objectives

- ◆ **Must be clear and definite objectives to be measured in order for evaluations to determine achievements**

Definition of Data to be Collected

- ◆ **What is to be measured in relation to the objectives must be determined**

Methodology

- ◆ **Evaluation design is made to allow for valid and reliable measurement**

Instrumentation

- ◆ **Data collection instruments are designed and pre-tested**

Data Collection

- ◆ **Process of collecting information that can be measured**

Data Processing

- ◆ **Data are put into a form that can be analyzed**

Data Analysis

- ◆ **Statistical tests are applied to the data to identify significant relationships**

Reporting

- ◆ **Evaluated results are compiled and reported**

Types of Data & Statistical Tests

- ◆ **Nominal**
- ◆ **Ordinal**
- ◆ **Interval Scale**
- ◆ **Ratio Scale**

Nominal:

- ◆ Variables are labels
- ◆ Name categories
- ◆ No numeric value

Examples

Married, urban, curly hair

- ◆ Statistical tests performed with nominal data:

Frequencies, Mode, Chi-Square

Ordinal:

- ◆ Variables imply intensity or severity.
- ◆ Categories are sequenced, ranked in order.

Examples

- ◆ First, second, third; youth, middle age, elderly
- ◆ Statistical tests performed with ordinal data:
 - Frequencies, mode, median

Interval Scale:

- ◆ Variables have a standard unit of measurement with no absolute zero

Examples

Temperature, standardized test scores

- ◆ Statistical tests performed with interval scale:

Frequencies, mean, T-test,
ANOVA/MANOVA, Pearson Correlation,
Regression Analysis

Ratio Scale:

- ◆ Variables have a standard unit of measurement with an absolute zero.

Examples

Weight, height

- ◆ Statistical tests performed with ratio scale:

Frequencies, mean, T-test,

ANOVA/MANOVA, Pearson

Correlation, Regression Analysis

Types of Reliability

- ◆ Test-Retest
- ◆ Alternative form method or multiple form
- ◆ Split half method
- ◆ Inter-observer or inter-rater reliability
- ◆ Intra-observer method or intra-rater reliability
- ◆ Internal consistency

Test-Retest

- ◆ Comparing results from one administration of an instrument with the results of a second administration of same instrument at a later time, using the same subjects.

Alternative form method or multiple form

- ◆ Compares the results of two forms of the same instrument

Split half method

- ◆ Divides an instrument in half and compares the results of one half against the results of the other.

Inter-observer or inter-rater reliability

- ◆ Compares the results obtained from one observer with the results of another observer using the exact same method.

Intra-observer method or intra-rater reliability

- ◆ Compares the results obtained by the same observer on the same subjects but at different times.

Internal consistency

- ◆ Measures the extent to which the items in the instrument are similar or measure the same concept.

Types of Validity

- ◆ Construct validity
- ◆ Content validity
- ◆ Internal validity
- ◆ External validity

Construct Validity

- ◆ Concerned with accuracy of the construct or concept that an instrument is attempting to measure

- **Example**

- ◆ If measuring 'locus of control' will need to find instrument with proven accuracy in measuring this.

Content Validity

- ◆ Concerned with the subjective determination of validity
- ◆ Uses some form of expert judgment
 - Examples
 - ◆ Literature reviews
 - ◆ Evaluation from experts in the field

Internal Validity

- ◆ If conclusions drawn correctly describe what happened in the study
- ◆ Degree of certainty that a program caused a change that is being measured
- ◆ Requires variables that are logically consistent that represent a testable causal relationship

External Validity

- ◆ Whether or not findings can be generalized
- ◆ Concerned with extent to which conclusions drawn from the research or evaluation can be applied to similar settings or populations outside the present study