The Retrospective Pretest Method for Evaluating Training

The slides and recording for this webinar are available from www.evalu-ate.org/webinars/2015-dec/

Overview

In a retrospective pretest, respondents rate themselves before and after an intervention in a single data collection event. It is useful for assessing individual-level changes in knowledge, skill, attitude, and behavior as one part of a overall evaluation of an intervention. This method fits well with the Kirkpatrick Model for training evaluation—see **www.kirkpatrickpartners.com** to learn more about that framework.

The main benefits of using this method are that it reduces response-shift bias. It is also convenient, more accurate than self-reported data gathered using traditional pre-post methods, adaptable to a wide range of contexts, and generally more acceptable to adult learners than traditional testing. Theodore Lamb provides an succinct overview of this method in the Harvard Family Research Project newsletter—see **bit.ly/hfrp-retro**.

The University of Wisconsin Extension's Evaluation Tip Sheet 27: Using the Retrospective Post-then-Pre Design provides practical guidelines about how to use this method: **bit.ly/uwe-tips**.

For a summary of research on this method, see Klatt and Powell's (2005) white paper, "Synthesis of Literature Relative to the Retrospect Pretest Design:" **bit.ly/retro-syn**.

Design

The focus of retrospective pretest questions should be on the knowledge, skills, attitudes, or behavior that are the focus of the intervention being evaluated.

General guidelines for formatting questions: 1) Use between 4 and 7 response categories in a Likert-type or partially anchored rating scale; 2) Use formatting to distinguish *pre* and *post* items; 3) Provide clear instructions to respondents. If you are using an online survey platform, check your question type options before committing to a particular format. To learn more about question formatting, see University of Wisconsin Evaluation Tip Sheet 28: "Designing a Retrospective Post-then-Pre Question" at **bit.ly/uwe-tips**.

For several examples of Likert-type rating scales, see **bit.ly/likert-scales**—be careful to match question prompts to rating scales.

Analysis & Visualization

Retrospective pretest data are usually ordinal, meaning the ratings are hierarchical, but the distance between the points on the scale (e.g., between "somewhat skilled" and "very skilled") are not necessarily equal. Begin you analysis by creating and examining the frequency distributions for both the pre and post ratings (i.e., the number and percentage of respondents who answer in each category). It is also helpful to calculate change scores—the difference between each respondent's before and after ratings—and look at those frequency distributions (i.e., the number and percentage of respondents who reported no change, reported a change of 1 level, 2 levels, etc.).

For more on how to analyze retrospective pretest data and ordinal data in general, see the University of Wisconsin Extension' Evaluation Tip Sheet 30: "Analysis of Retrospective Post-then-Pre Data" and Tip Sheet 15: "Don't Average Words" **bit.ly/uwe-tips**.

For practical guidance on creating attractive, effective bar, column, and dot plot charts, as well as other types of data visualizations, visit **stephanieevergreen.com**.

Using Results

To use the results of retrospective pretest questions to make improvements to an intervention, examine the data to determine if some groups (based on characteristic such as job, other demographic characteristics, and incoming skill level) gained more or less than others and compare results to the intervention's relative strengths and weaknesses in terms of achieving its objectives. Make adjustments to future offerings based on lessons learned and monitor to see if the changes lead to improvements in outcomes.

This material is based upon work supported by the National Science Foundation under grant number 1204683. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the presenters and do not necessarily reflect the views of NSF.