

Application of Crossover Design for Conducting Rigorous Extension Evaluations

Abstract

With the increasing demand for accountability of Extension programming, Extension professionals need to apply rigorous evaluation designs. Randomized designs are useful to eliminate selection biases of program participants and to improve the accuracy of evaluation. However, randomized control designs are not practical to apply in Extension program evaluation. This article explains how to use the crossover design as a practical tool for evaluating Extension programs rigorously. This design can be used to evaluate any Extension program with two or more curricula presented to client groups in multiple counties.

K. S. U. Jayaratne
State Leader for
Program Evaluation
and Associate
Professor
Department of
Agricultural and
Extension Education
jay_jayaratne@ncsu.edu

Carolyn L. Bird
Associate Professor
and Family Resource
Management
Specialist
Department of 4-H
Youth Development
and Family &
Consumer Sciences
carolyn_bird@ncsu.edu

**Jacquelyn W.
McClelland**
Professor, Nutrition
Specialist
Department of 4-H
Youth Development
and Family &
Consumer Sciences
jackie_mcclelland@ncsu.edu

North Carolina State
University
Raleigh, North
Carolina

Introduction

Extension stakeholders demand systematic evaluations that document program outcomes (Bailey & Deen, 2002; Stup, 2003). Therefore, agents and specialists are compelled to evaluate programs rigorously. However, some evaluations still are not rigorous due to lack of evaluation capacity (Chapman-Novakofski, et al., 1997), tools, and methods. Retrospective pre- and post-test design (Davis, 2003) is the most commonly practiced evaluation in Extension.

One advantage of randomized control design is it minimizes the threats to the internal validity by eliminating selection bias. Randomization "helps to distribute the idiosyncratic characteristics of participants over the treatment levels so that they do not selectively bias the outcome of the experiment" (Kirk, 2009, p. 24). Extension professionals need to know how to apply randomized control design to evaluate programs.

The crossover method is a randomized control experimental design in which treatments are switched for participants in such a way that everyone will be exposed to all the treatments, just in different time periods over the course of the experiment (Bate & Jones, 2006). Therefore, a crossover design

is practically and ethically appropriate to use in Extension evaluation. A crossover design may be uniform on the number of subjects, the number of periods, or uniform on both (Bate & Jones, 2006).

A drawback of crossover design is that one treatment may have residual effects and alter the response to subsequent treatments (Sibbald & Roberts, 1998). However, it is a common practice to consider carry-over effects to be first-order (carrying over to the next period only) (Bates & Jones, 2006). Uniform crossover designs have been widely used in medical and clinical trials, agricultural animal feeding trials, and other applications (Bate & Jones, 2008). Thus uniform crossover designs have proven to be an efficient and effective method of testing treatment effects.

Purpose

The purpose of this article is to explain the crossover design for evaluation of a program with two training curricula.

Evaluated Program

The Extension program evaluated consisted of two curricula, Eat Smart, Stay Well (ESSW) and Eating Well on a Budget (EWOB). The ESSW included a healthy diet, effects of dietary fats, and benefits of a diet high in fruits and vegetables. The EWOB focused on limited food dollar management for nutritious foods.

Participating Extension agents were trained to present and evaluate both curricula. Each of the agents had one group of participants. The program was presented in 10 sessions over 10 weeks. Five sessions were for ESSW, and five sessions were for EWOB. The results of the ESSW program are presented in "Nutrition Education Brings Behavior and Knowledge Change in Limited-Resource Older Adults" (McClelland, Jayaratne, & Bird, 2013).

Evaluation Design

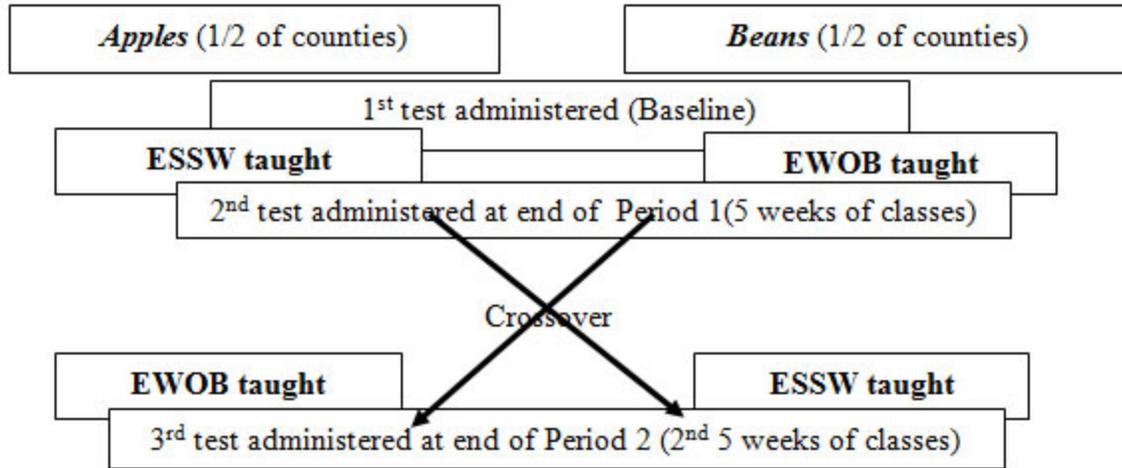
The survey consisted of 10 knowledge-testing and four behavior-testing questions, with half of each type of question drawn from each curriculum for equal representation. The survey was administered to test participants' knowledge and behavior at baseline (1st test), after five weeks (2nd test) and after 10 weeks (3rd test); each test was identical in content.

Knowledge-testing questions were true and false answer format. Behavior recording questions were five-point Likert scale and ranged from 1 = not practicing the behavior to 5 = practicing it regularly. Half of the counties were randomly selected to have the ESSW curriculum first (this group was named *Apples*); the other half received EWOB curriculum first (this group was named *Beans*).

Before beginning lessons, the first test was administered to both groups to assess participants' baseline knowledge and behavior related to ESSW and EWOB. Then, the *Apples* participants started with the ESSW curriculum, and the *Beans* participants started with the EWOB curriculum and continued for 5 weeks. At the end of the 5-week time period, the second test was given to each group. The second test documented participants' knowledge and behavior after their exposure to their respective curriculum.

Next, the curricula were switched for the groups. So the *Apples* received the EWOB, and the *Beans* received the ESSW. At the end of the second round of 5 weeks, the third test was administered to all groups to document their knowledge and behavior related to ESSW and EWOB curricula (Figure 1).

Figure 1.
Evaluation Design



Conducting Evaluation

The organization and interpretation of results are demonstrated in Table 1. The first five knowledge testing questions and two behavior testing questions were used to assess changes related to ESSW curriculum objectives.

For Period 1, the difference between the second test and the baseline results of the ESSW questions for the *Apples* provides the outcome data of the ESSW curriculum. The difference between the second test and the baseline results of the ESSW questions for the *Beans* provides the comparison data for the ESSW curriculum, because the *Beans* were not exposed to the ESSW curriculum before taking the second test.

Similar to this, the difference between the second test and the baseline results of the EWOB questions for the *Beans* provides the outcome data of the EWOB curriculum. The difference between the second test and the baseline results of the EWOB questions for the *Apples* provides the comparison data for the EWOB curriculum. The *Apples* were not exposed to EWOB curriculum before taking the second test.

For Period 2, the *Beans* received the ESSW curriculum, and the *Apples* received the EWOB curriculum. Then the third test was conducted with the two groups. The difference between the third test and the second test results of the ESSW questions for the *Beans* provides the outcome data of the ESSW curriculum in the second round. This can be considered a replication of the treatment. The difference between the third test and the second test results of the ESSW questions for the *Apples* provides the comparison data for the ESSW curriculum.

Similar to this, the difference between the third and the second test results of the EWOB questions for the *Apples* provides the outcome data of EWOB curriculum in round two. The difference between the third test and the second test results of the EWOB questions for the *Beans* provides the comparison data for the EWOB curriculum.

Table 1.
Organization and Interpretation of Evaluation Data

Randomized Groups	Knowledge and Behavior Testing Questions	<u>First Period</u> (2nd test) - (Baseline) (5 weeks)	<u>Second Period</u> (3rd test) - (2nd test) (10weeks)
<p><i>Apples</i> Received ESSW curriculum first. After 2nd test, this group received EWOB curriculum.</p>	5 knowledge testing questions for ESSW	Outcome data for ESSW curriculum (ESSW Treatment data)	Comparison data for the 2nd Period of ESSW curriculum (ESSW Control data)
	2 behavior testing questions for ESSW		
	5 knowledge testing questions for EWOB	Comparison data for 1st Period of EWOB curriculum (EWOB Control data)	Outcome data for EWOB curriculum (EWOB Treatment data)
	2 behavior testing questions for EWOB		
<p><i>Beans</i> Received EWOB curriculum first. After 2nd test, this group received ESSW curriculum.</p>	5 knowledge testing questions for ESSW	Comparison data for the 1st Period of ESSW curriculum (ESSW Control data)	Outcome data for ESSW curriculum (ESSW Treatment data)
	2 behavior testing questions for ESSW		
	5 knowledge testing questions for EWOB	Outcome data for EWOB curriculum (EWOB Treatment data)	Comparison data for the 2nd Period of EWOB curriculum (EWOB Control data)

	2 behavior testing questions for EWOB	
--	------------------------------------------------	--

Steps for Analyzing the Outcomes of ESSW Curriculum

1. Conduct independent sample t-test to determine whether the treatment and control groups are comparable for ESSW knowledge test scores/behavior test scores at baseline.
2. Compare means of knowledge test scores/behavior test scores of baseline data to test 2 data (after completing the 1st period for ESSW treatment) using paired sample t-test to determine whether there was an effect of the curriculum.
3. Compare means of knowledge test scores/behavior scores of baseline data to test 2 data (after completing the 1st period for ESSW control) using paired sample t-test to determine whether there wasn't any significant effect.
4. Complete the steps 2 and 3 for the 2nd Period with ESSW treatment and control group data to determine whether the replicated trial has similar results.
5. Compare mean differences of knowledge test score/behavior test score for before and after completing the ESSW curriculum for treatment and control groups to determine whether the curriculum has significant improvement over the treatment.

Follow these steps to determine EWOB outcomes.

Practical Implications

Crossover design can be applied for evaluating any Extension program with two or more identifiable curricula. It provides a practical option for Extension professionals to conduct rigorous evaluations of programs.

References

Bailey, S. J., & Deen, M. Y. (2002). A framework for introducing program evaluation to extension faculty and staff. *Journal of Extension* [On-line], 40(2) Article 2IAW1. Available at: <http://www.joe.org/joe/2002april/iw1.php>

Bate, S. T., & Jones, B. (2006). The construction of nearly balanced and nearly strongly balanced uniform cross-over designs. *Journal of Statistical Planning and Inference* , 136(9), 3248-3267.

Bate, S. T., & Jones, B. (2008). A review of uniform cross-over designs. *Journal of Statistical Planning and Inference* , 138(2), 336-351.

Chapman-Novakofski, K., Boeckner, L. S., Canton, R., Clark, C. D., Keim, K., Britten, P., & McClelland, J. (1997). Evaluating evaluation—What we've learned. *Journal of Extension* [On-line],

35(1) Article 1RIB2. Available at: <http://www.joe.org/joe/1997february/rb2.php>

Davis, G. A. (2003). Using a retrospective pre-post questionnaire to determine program impact. *Journal of Extension* [On-line], 41(4) Article 4TOT4. Available at: <http://www.joe.org/joe/2003august/tt4.php>

Kirk, R. E. (2009). Experimental design. In R. E. Millsap and A. Maydeu-Olivares (Eds.), *The SAGE handbook of quantitative methods in psychology*. Thousand Oaks, California: SAGE Publications Inc.

McClelland, J., Jayaratne, K. S. U., & Bird, C. (2013). Nutrition education brings behavior and knowledge change in limited-resource older adults. *Journal of Extension* [On-line], 52(2) Article 2FEA1. Available at: <http://www.joe.org/joe/2013april/a1.php>

Sibbald, B., & Roberts, C. (1998). Understanding controlled trials: Crossover trials. *BMJ*. 316(1719). Retrieved from: <http://www.bmj.com/content/316/7146/1719.full>

Stup, R. (2008). Program evaluation: Use it to demonstrate value to potential clients. *Journal of Extension* [On-line], 41(4) Article 4COM1. Available at: <http://www.joe.org/joe/2003august/comm1.php>

Copyright © by *Extension Journal, Inc.* ISSN 1077-5315. Articles appearing in the Journal become the property of the Journal. Single copies of articles may be reproduced in electronic or print form for use in educational or training activities. Inclusion of articles in other publications, electronic sources, or systematic large-scale distribution may be done only with prior electronic or written permission of the Journal Editorial Office, joe-ed@joe.org.

If you have difficulties viewing or printing this page, please contact [JOE Technical Support](#)