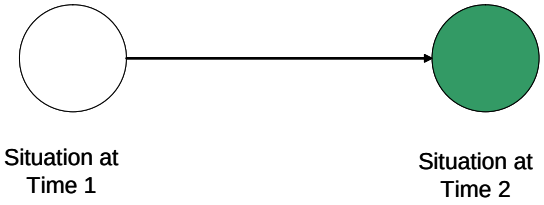
Causality, Attribution and USAID Evaluations

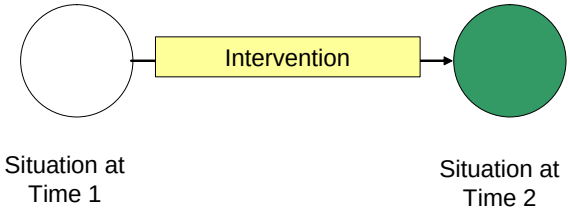
We are used to presuming that if something changes – a situation or condition – there must be some explanation, some reason for the change:



Depending on the situation or condition we are looking at we may be inclined to think that:

- Time alone accounts for the change, e.g., baby starts speaking (maturation)
- Any of a number of causes might have produced the change, e.g., higher income.
- There are a limited number of likely causes of the change, and a high probability that one specific cause produced the change, e.g., lung cancer.
- There is one and only one cause that plausibly explains the change and in the absence of that particular cause, the change would not have occurred.

Sometimes we know that something intervened between Time 1 and Time 2 – a policy change or the provision of some new service -- and that over the same time period the condition or situation changed. But two things being true do not prove that one caused the other. Correlation does not equal causality.



If we want to say with confidence that the intervention caused the change, we still have to produce evidence that shows that:

- The intervention actually caused the change, i.e., nothing would have changed in the absence of the intervention. The intervention was ***necessary***.
- The intervention was the only cause of the change, nothing else was needed to bring about the change. The intervention was ***sufficient***.

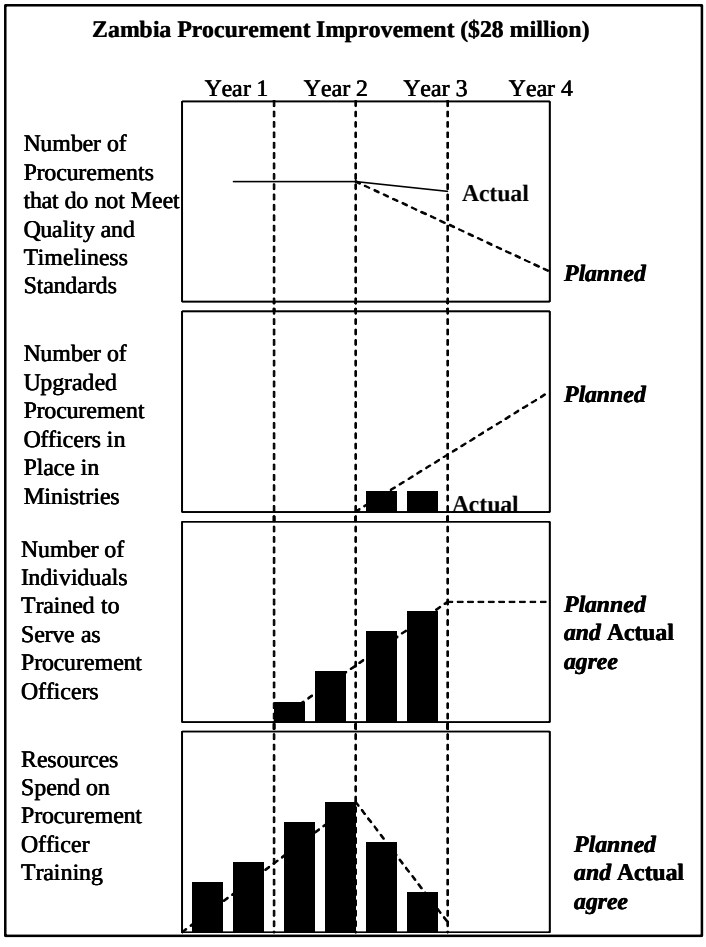
When we attribute a change in a situation or condition to a USAID project, we are saying that we have sufficient evidence to demonstrate that the USAID project was both ***necessary and sufficient*** to cause the change.

OMB, the GAO and other observers of U.S. government projects and programs, including those USAID funds, have becoming increasingly concerned about the adequacy of the evidence used to support claims that changes in situations and conditions are attributable to U.S. government funded programs and projects.

Good performance monitoring systems that focus only on program beneficiaries as well as on the activities carried out by project implementers often show what changes occurred at both the Output and Outcome level.

Performance Indicators	Year 1		Year 2		Year 3	
	Planned	Actual	Planned	Actual	Planned	Actual
Percentage of Rural Families Living in Poverty	38%	38%	30%	32%	25%	27%
Percentage of Farmers Using New Seeds and Technology	20%	22%	40%	38%	60%	56%

When organized and analyzed by result levels, performance monitoring data can also suggest whether development hypotheses have proven to be valid or whether there were breaks in the cause-and-effect results chain, as was the case in the Zambia project illustrated here. A vertical analysis of monitoring data will quickly highlight problems, but positive results at all levels of a hierarchy of results may not be sufficient to prove that Outputs caused Outcomes.



A. Using Evaluations to Prove Cause and Effect Relationships

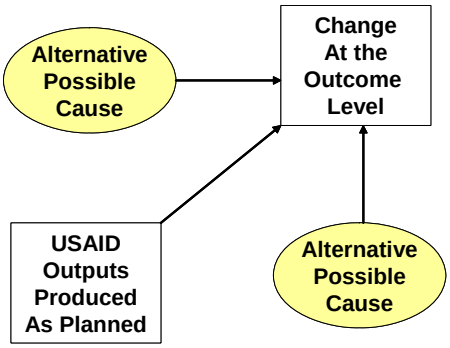
Evaluations help generate evidence about causality by eliminating other possible explanations (causes) for documented changes at the Outcome level.

After the fact approaches include:

1. After the Fact Approaches

Evaluations at the end of a project cannot go backwards to strengthen the case for attribution, they have to start with what is available.

a. Elimination of Other Plausible Explanations



An evaluation can help to provide a basis for attributing changes at the Outcome level to USAID by gathering data on alternative possible causes and whether and to what degree they may have played a role. Likely candidates in many environments are government policies and services (which may have changed) and other donor programs. If the evaluation can demonstrate other programs, activities or policy shifts did not occur or did not affect the particular region or target group, the case for attribution to USAID is strengthened, but may still not be conclusive. MSI used this methodology in a series of "rule of law" assessments in Eastern Europe calling it a "back pass" methodology, working from results backwards to causes. At nearly the same time, UNDP introduced what it calls "outcome" evaluations, which use the same technique.

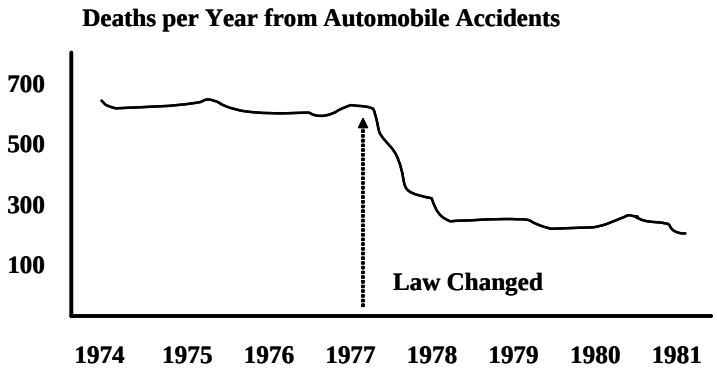
b. Post-Facto Comparisons

	Before (Baseline)	After Intervention
Area or Population Was Affected By the Intervention	38%	27%
Area or Population Was Not Affected By the Intervention		32%

Even when there was no effort made at the start of a project to set up comparisons that would help to prove whether a project intervention supplied the necessary and sufficient conditions required to produce changes in important Outcomes, it is sometimes possible to demonstrate that in areas and target groups that were not affected by the project, the status of conditions at the Output level are less desirable or positive than in areas or with target groups affected by the project. The key problem with evaluations that use this approach to producing evidence that support assertions of causality and attribution is that the status of situations chose as comparisons prior to the intervention period is often not know. It may be that those areas or populations changed just as much as did the target areas or populations.

c. Mixed Approach

In some instances, where everyone has been affected by a change, e.g., new laws, evaluations may be able to find time series data that can be used to determine whether a project intervened in a way that interrupted a time series that closely measures the Outcome the project was designed to affect. An interrupted times series design may need to be used in



combination with an effort to demonstrate that nothing else was going on at the time (no alternative cause) that could explain the change in the time series. The classic example in this case was Donald Campbell's use of highway accident figures to demonstrate that a change (reduction) in the speed limit in the Connecticut lowered the automobile death rate.

2. Before the Fact Approaches

When the need for evidence about causality or to support attribution of changes to a USAID project is significant[1] and is recognized before or as a project is being developed, evaluation designs can be put in place that will significantly improve the quality of evidence about causality. Most designs of this sort insert comparison groups into the picture at the start of the project and track changes in both the program and control groups to determine the impact of a project. .

	Before (Baseline)	After Intervention
Area or Population Was Affected By the Intervention	38%	27%
Area or Population Was Not Affected By the Intervention	38%	32%

Conceptually, and even from a cost point of view, it is not that hard to add this kind of comparison to the design for a projects where being able to prove the extent to the projects effect is important. What turns out to be difficult is selecting the comparison or "control" areas or groups. For comparisons to be valid, the areas or groups must be as similar to those that are affected by the project intervention as possible. The importance of being able to demonstrate that the populations or areas that were and were not affected by a project intervention are virtually identical is what has led OMB and the Poverty Action Laboratory at MIT to encourage the "randomized assignment" of well matched people or areas to the project group that will be affected by the intervention and the "control" group that will not.[2] Approaches of this sort can also help managers to determine what components of programs have the greatest influence on results, as a quasi-experimental evaluation design for the DevTech managed USAID/Malawi's Safe Schools program, along the lines shown below, illustrates.

Group A	Group B	Group C	Group D
Only students are trained	Only teachers are trained	Both students and teachers are trained	No training is provided (control)

Use of these more sophisticated evaluation designs is increasing in USAID, as the attached examples indicate. They are not right for every situation or even most projects and programs. But they are appropriate when strong evidence of causality is needed.

Drafted: MHageboeck, MSI, 9/3/07

---

1 The need for strong evidence of causality is often most significant for projects where the intent, from the start, is to test a model with the idea that it will be replicated or scaled up.

2 OMB's guidance: http://www.whitehouse.gov/omb/part/2004_program_eval.pdf

A Quasi-Experimental Study to Assess the Performance of a Reproductive Health Franchise in Nepal

By Sohail Agha, Ali Mehryar Karim, Asma Balal, Steve Sossler

Contributed By CMS Project

Download Now

Published: 2003

In 2001, the Commercial Market Strategies project established a nurse and paramedic franchise in Nepal to increase utilization of reproductive health services and client satisfaction with service quality. To assess the impact of the intervention, CMS used a quasi-experimental study design, with baseline and follow-up measurements on non-equivalent control groups. CMS found that at the clinic level, client satisfaction increased at intervention but not at control clinics. Client loyalty, measured by return visits, also increased at intervention but not at control clinics. The increase in client loyalty was, in part, explained by the increase in satisfaction with service quality. At the population level, CMS did not find consistent increases in utilization of various reproductive health services. While utilization of other reproductive health services did not change, an increase in contraceptive use may have been associated with use of the nurse and paramedic network. CMS concluded that a franchiser that provides training to franchised clinics in reproductive health service delivery and in client-provider interaction, and that monitors the quality of care provided at these clinics, can help increase client satisfaction at network clinics.

Impact Evaluation of Active Labor Programs--Romania

The overall goal of this project was to provide information on the net impact of selected active labor market programs (ALMPs) in Romania and to improve the targeting and cost-effectiveness of these programs. Specifically, we analyzed the net impacts of four Romanian ALMPs:

- Training and Retraining (TR)
- Small Business Consultancy and Assistance (SB)
- Public Works Community Job Creation (PW)
- Employment and Relocation (ER)

To measure the net impacts of these programs, we used a rigorous quasi-experimental evaluation technique. We selected a representative sample of program participants and non-participants from 15 judets throughout Romania. We interviewed members of these two groups, completing a total of 3,999 interviews (2,050 participants and 1,949 non-participants). Using the propensity score matching technique, for each of the four ALMPs, we identified matched participant and comparison samples. Data from these matched samples were then used to derive program impact estimates.

Our analysis of program impacts revealed that three of the four programs (TR, SB and ER) had success in improving participants' economic outcomes. Thus, the impact estimates from this evaluation indicate that ALMPs in Romania had mostly positive impacts on employment and earnings. These results, however, were not uniform for all programs and for all subgroups. As a result, the effectiveness of ALMPs in Romania can be greatly enhanced by using these results to target services in the future.

PROJECT
India: Livelihood Skills and Opportunities for Girls in Allahabad

In India, 40 percent of girls ages 15–19 are married. Livelihood activities could help reframe the second decade of girls' lives from a period when they are often confined to the home and devoted to preparation for marriage and childbearing to a time when they can develop as individuals and gain knowledge and skills that are the foundation for a more productive adulthood. Moreover, by offering an alternative source of social status, work is also likely to delay marriage and provide women with greater control over their sexual and reproductive lives.

The Population Council, in collaboration with CARE India, has tested the feasibility and impact of adding four additional components to a reproductive health project in urban slum areas of Allahabad, Uttar Pradesh:

- counseling about savings formation and livelihoods;
- training in vocational skills;
- assistance in opening savings accounts; and
- follow-up support

Using a quasi-experimental pre- and post-test design that contrasted the experimental group with a comparison group of adolescents, the project investigated whether the intervention:

- increased girls' physical mobility and contact with individuals outside the family;
- enhanced girls' skills development and sustained use of these skills;
- altered work aspirations of girls and encouraged more progressive gender role norms;
- reduced gender differentials in time use; and
- increased girls' reproductive health knowledge

The baseline survey, conducted in May 2001 among over 3,000 adolescents ages 14–19 and their parents, revealed substantial differences between boys and girls with respect to mobility, time-use patterns, savings and work experience, and reproductive health knowledge. Data collection for the project's endline survey began in March 2003 and was completed in July 2003. Data entry and cleaning were completed in October 2003. The analysis showed that participation of unmarried adolescent girls increased in all project activities. Vocational counseling, follow-up support for unmarried adolescent girls, and encouragement to form group savings schemes were included in a subsequent CARE project, Action for Slum Dwellers' Reproductive Health.

Analysis of the endline data in combination with the baseline data indicated that although the livelihoods program was acceptable to parents and feasible to implement, the project had only a minimal impact on the behavior and attitudes of adolescent girls in the experimental slums. The greatest changes between the baseline and the endline surveys were found in those outcomes that most closely reflected the content of the intervention. Girls in the intervention group were significantly more likely to have knowledge of safe spaces, be a member of a group, score higher on the social skills index, be informed about reproductive health, and spend time on leisure activities than the matched control respondents. No effect was found on gender role attitudes, mobility, self-esteem, work expectations, time use, or labor market work likely because of the short duration of the intervention, as well as the limited number of times that groups convened. Those designing future livelihoods interventions with adolescent girls are advised to extend the period of time spent on group formation, negotiation and social skills, and vocational skill development.

Location: Allahabad, Uttar Pradesh, India Duration: 2001–2005

Population Council researchers

Barbara S. Mensch, Monica J. Grant, Paul C. Hewett, Mary Philip Sebastian

A quasi-experimental study to assess the impact of four adolescent sexual health interventions in Sub-Saharan Africa

International Family Planning Perspectives, Jun 2002 by Agha, Sohail

CONTEXT: Rigorous evaluations are needed to assess whether adolescent sexual health interventions have an effect on young people's risk-related perceptions and behaviors.

METHODS. A quasi-experimental design was used to evaluate the impact of adolescent sexual health interventions conducted by social marketing programs in Cameroon, Botswana, South Africa and Guinea in 1994-1998. The some statistical models, using data from baseline and postintervention surveys, were employed to study each intervention; the results are presented within the framework of the Health Belief Model.

RESULTS. The interventions were associated with improvements in a variety of health perceptions among women, including perceptions of benefits of and barriers to protective behavior; for women, the interventions also had positive impacts on contraceptive use. Effects were much more limited among men, although evidence from Cameroon and Botswana suggests that men were less likely after the intervention than before to have multiple or casual partners. The Cameroon intervention, the most successful of the four, used multiple communications media (including radio and peer education) and reached nine in 10 adolescents, the Botswana program also reached a high proportion of the target audience. In South Africa and Guinea, however, the programs were less intensive and had a more limited reach. CONCLUSIONS: Interventions targeted at adolescents can be effective in changing attitudes and sexual behavior if they include multiple channels of communication, reach a substantial proportion ofyoung adults and make contraceptives widely available. There remains an urgent need to identify ways to address young men's sexual health concerns effectively.