There are no pending text suggestions from your readers

Selecting impact/outcome evaluation designs: a decision-making table and checklist approach

A topic article in the Outcomes Theory Knowledge Base

Once it has been decided that using an impact/outcome evaluation design is likely to be appropriate, feasible and affordable in regard to a particular intervention, the next question is working out which is the most appropriate impact/outcome design to use. Making this selection can be complex, particularly for those not familiar with all of the possible options. This article provides a decision-making table approach to selecting one of seven possible groups of impact/outcome evaluation designs. It also provides a set of checklists for each of the seven possible types of impact/outcome evaluation. In addition, it helps in selecting the best technique for attempting to deal with non-equivalent comparison group issues in the pragmatic, and hence frequently used, constructed matched comparison group designs.

outcomes theory
outcomes theory knowledge base

Introduction to outcomes theory
List of outcomes theory articles
Subscribe to author's e-newsletter
Author's Outcomes Blog
Follow author on Twitter

Introduction

[Please note: the decision-making table and checklist approach set out in this article is in its initial iteration, therefore any comments on ways of

improving it would be much appreciated, please leave any comments at the bottom of this page]

Deciding on an impact/outcome evaluation design can be difficult due to the fact that there is a large range of options for the way that such evaluations can be designed. This article uses a set of decision-making tables to assist in deciding on the the most appropriate impact/outcome design to use. It then also looks at how to select techniques to deal with potentially non-equivalent comparison groups in the pragmatic, and hence popular, constructed matched comparison group designs. Lastly, it provides seven separate checklists, one for each of the seven possible impact/outcome evaluation designs which are used in the tables. These checklists can be used to assist in impact/evaluation design selection decisions and also when peer reviewing an impact/evaluation design which is being proposed or has been carried out. (More information on the seven designs is available at: Seven possible impact/outcome evaluation designs. An article which takes a very simple example and uses it to illustrate what each of the impact evaluation designs and techniques described in this article look like in practice is available at: Impact/outcome evaluation designs and techniques illustrated with a simple example.

The decision to use an impact/outcome evaluation design

It should never be a assumed that an impact/outcome evaluation design is always the most appropriate design to use in an evaluation. There are three different types of evaluation - *formative evaluation* which attempts to help optimize intervention implementation; *process evaluation* which describes the course and context of an intervention; and *impact/outcome evaluation* which attempts to attribute changes in high-level outcomes to an intervention. It is often much more strategic to set up an evaluation which uses limited formative evaluation to ensure that an intervention is implemented in an optimal manner; uses limited process evaluation to identify best practice; but which does not use impact/outcome evaluation. This is because impact/outcome evaluation is often costly, difficult and hard to ensure that it is done properly. Impact/outcome evaluation resources are often better pooled and impact/outcome evaluations undertaken on selected pilot programs where evaluators have sufficient resources and control over the intervention and its context to maximize the chances of a robust and successfully completed impact/outcome evaluation. The findings from these pilots can then be used in the full roll-out phase of programs by ensuring that best practice is applied (but often without doing an impact/outcome evaluation) (see: Impact evaluation - when it should and should not be done; Full roll-out impact/outcome evaluation versus piloting impact/outcome evaluation plus best practice monitoring; Best practice representation and dissemination using outcomes models).

Impact/outcome evaluation design can be a complex technical decision

However, in those instances where it has been decided that an impact/outcome evaluation design is going to be used in an evaluation, deciding on which design to use can be a highly technical exercise. The evaluation designer is expected to be familiar with all of the potential impact/outcome evaluation designs. The purpose of the decision-making table approach used in this article is to provide a framework which evaluation designers can work with over time as a basis for thinking about impact/outcome evaluation designs and for justifying why they have made a particular design decision. In the first instance, an evaluation designer who is not very familiar with impact/outcome evaluation designs will almost certainly have to talk through the use

of the tables in this article with someone who has more experience. However if they work with this framework, over time it should be possible for them, through the repeated use of the decision-making tables in this article, to develop a more consistent and robust framework for their thinking about evaluation design. Advanced impact/outcome evaluation designers, will want to be more creative in their designs and will be able to develop hybrid designs (e.g. randomized experiments combined with other elements) which are closely tailored to the complexities of the particular intervention they are evaluating and its context.

How to use this decision-making table approach

Below are two decision-making tables. Table 1 is the *Impact/outcome evaluation design selection table*. To use this table, look along the top of the table and identify those things which apply to the intervention you are wanting to evaluate. If one of them applies to the particular intervention you are looking at, then look down the column immediately below it. Then look across to the left-hand side of the table and you will see which of the seven possible types of impact/outcome evaluation may be appropriate in your particular situation.

Table 2 *Selecting techniques to improve constructed matched comparison group designs* applies to situations where there may be doubts about a comparison group being equivalent to an intervention group (i.e. in constructed matched comparison group designs). Within this table look down the column on the left-hand side of the table to identify those things which apply to the intervention for which you are designing the evaluation. Then select a technique which is appropriate for your particular intervention.

Preliminary information

The following section sets out preliminary information on what is contained in the decision-making tables to make it easier for the user to better understand the tables.

Seven types of impact/evaluation design

The decision-making tables are based on a framework which divides impact/outcome evaluation designs into seven possible types. These are:

1. True experimental design
2. Regression-discontinuity design
3. Time-series analysis
4. Constructed matched comparison group design
5. Exhaustive causal identification and elimination design
6. Expert judgment design

7. Key information judgment design

Note that not all of these designs (particularly the last two) are accepted by all stakeholders as robust impact/evaluation designs, but all are accepted by some stakeholders in some situations as appropriate, feasible and affordable and hence are included for the sake of completeness.

(More information is available on these designs at [Seven possible impact/outcome evaluation designs](#)).

Decision-making Table 1: Impact/outcome evaluation design selection

Decision-making Table 1 focuses on the list of seven possible impact/outcome evaluation designs running from *true experiments* through to *expert judgment designs*. Working out which impact/outcome designs are the most appropriate can be assisted by looking at the features of the intervention being considered for impact/outcome evaluation. Interventions may have one or more of the following features (which are listed along the top of Decision-making Table 1):

- **Cannot control who gets intervention.** Where evaluators cannot control who gets an intervention, this seriously limits the range of possible impact/outcome evaluation designs they can use.
- **Cannot stop control/comparison group getting any intervention.** In some cases, evaluators cannot stop a control or comparison group (if it is being used in an impact/outcome evaluation design) from also getting a version of the intervention. This obviously causes problems for any evaluation design based on comparing outcomes for an intervention group versus a control or comparison group.
- **Intervention applied to all.** Where an intervention is applied to everyone (all people or units) this seriously limits the range of impact/outcome evaluation designs which can be used.
- **Causal mechanisms difficult for key informants/experts to discern**. Due to the nature of some interventions, it is relatively easy to observe the connections between the intervention and outcomes. However, in other cases it is possible that many things are influencing final outcomes. Different designs are more appropriate in different circumstances.
- **Limited resources for impact/outcome evaluation.** There are often limited resources for evaluation and, as a consequence, for impact/outcome evaluation. This can influence the type of impact/outcome evaluation design which is chosen.
- **Powerful stakeholders skeptical about intervention.** Where powerful stakeholders are skeptical about the effectiveness of an intervention, more intense forms of evaluation are often needed which such stakeholders will take as having 'proved' that an intervention has caused final outcomes to improve.
- **Concern about risk of impact/outcome evaluation not being completed.** It is easy to waste significant resources on impact/outcome evaluations which are never completed because evaluation planners have underestimated the practical problems in actually bribing an impact/outcome evaluation to completion. This risk needs to be factored into evaluation design.

Four techniques for improving constructed matched comparison impact/outcome designs

Of the seven impact/outcome evaluation designs, the constructed matched comparison group design is often chosen as a pragmatic design because it can be used in (the often occurring) cases where an evaluator does not have control over who gets an intervention. Such a lack of evaluator control rules out possibilities such as a true experiment or a regression-discontinuity design. However, constructed matched comparison group designs often suffer from the actual, or potential, problem that the comparison group differs in some ways from the intervention group. This problem can be reduced or dealt with by using one of four techniques. These techniques are:

1. Difference-in-difference
2. Instrumental variables
3. Propensity matching
4. Case matching

 (More information on these techniques is available at [Techniques for improving constructed matched comparison group impact/outcome evaluation designs](#).

Decision-making Table 2: Selection of techniques to improve constructed matched comparison group impact/outcome evaluation designs

Constructed matched comparison group designs are often considered (because they are usually more feasible than true experiments) and can often be done even where an evaluator does not have control over who receives an intervention. While such designs are often much more acceptable (e.g. from an ethical point of view) and feasible (e.g. from an intervention allocation point of view) a major problem often arises because there is a high risk that the comparison group which is used will differ in important ways from the intervention group. In a number of cases of constructed matched comparison group designs, the following features of the situation can mean that particular techniques can be used to improve, or deal with, differences between the comparison group and the intervention group. These features are set out in the left-hand column of Decision-making Table 2.

- **Can track trends in comparison and intervention group.** Where it is possible to track trends in both the comparison and intervention group.
- **Can find another characteristic (variable) causing people (or units) to not go into the intervention group, unrelated to outcome**.
  Such characteristics can be used to identify a subset of potential comparison group members who are more likely to be like intervention group members.

Selecting impact/outcome evaluation designs: a decision-making table and checklist approach – a knol by Paul Duignan, PhD

15/08/09 7:32 AM

- **Can describe the intervention group and potential comparison group members accurately so predictions can be made as to the likely outcome for someone (or some unit) in the absence of the intervention.** Good data on intervention group and potential comparison group members can be utilized with mathematical techniques to predict the expected outcomes for someone not receiving the intervention.
- **Can construct a comparison group by locating other individuals (or units) which are exactly 'matched' with members of the intervention group on key characteristics (variables).** This technique can be used to ensure that the members of the intervention and comparison group are sufficiently similar.

The decision-making tables

Below are the two decision-making tables. The first table should be used by looking along the top line and selecting the features which apply to the intervention being considered for impact/outcome evaluation. Then the user should look down the appropriate column and at the list of impact/outcome evaluation designs on the left to determine which design is most likely to be the appropriate. (More information on the Seven possible impact/outcome evaluation designs is available).

Contents

Impact/outcome evaluation design selection

	☐ Cannot control who gets intervention?	☐ Cannot stop control/comparison group getting any intervention?	☐ Intervention applied to all?	☐ Causal mechanisms difficult for key informants/ experts to discern?	☐ Limited resources for impact/outcome evaluation?	☐ Powerful stakeholders skeptical about intervention?	☐ Concern about risk of impact/outcome evaluation not being able to be completed?
☐ True experiment◢	Not appropriate	Not appropriate	Not appropriate	May be appropriate	May be more expensive	May be more appropriate	Higher
☐ Regression discontinuity ◢	Not appropriate	Not appropriate	Not appropriate	May be appropriate	May be more expensive	May be more appropriate	Higher
☐ Time series analysis ◢	May be appropriate	Does not rely on a control group	May be appropriate	May be appropriate	May be less expensive	May be more appropriate	Lower
☐ Constructed matched comparison group◢	May be appropriate	Not appropriate	Not appropriate (except different time period)	May be appropriate	May be less expensive	May be more appropriate	Lower
☐ Exhaustive causal identification and elimination ◢	May be appropriate	No formal control group	May be appropriate	Not appropriate	May be less expensive	May be less appropriate	Lower
☐ Expert judgment *◢	May be appropriate	No formal control group	May be appropriate	Not appropriate	May be less expensive	May be less appropriate	Lower
☐ Key informant *◢	May be appropriate	No formal control group	May be appropriate	Not appropriate	May be less expensive	May be less appropriate	Lower

* Particular stakeholders may reject the last two designs, or some of the other designs above, as being inadequate to robustly establish causality in specific cases, or in all cases.
Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D135-0.
Comments for improving this decision-making table welcome - please post at the end of the article referred to immediately above.

doview.com model

Table 1: Impact/outcome evaluation design selection

To use the second table, which particularly applies to constructed matched comparison group designs, look down the column on the left-hand side, identify the features relevant to the intervention being considered for impact/outcome evaluation and select the appropriate technique.

Selecting techniques to improve constructed matched comparison group designs

What is possible in the situation	Technique	How to do it
Can track trends in comparison and intervention group? Comparison group starts off different from the intervention group but trends in the comparison group and intervention group can be tracked over time.	Difference-in-difference	Track both intervention and comparison group and work out what the improvement in the intervention group is OVER AND ABOVE any improvement occurring in the comparison group.
Can find another characteristic (variable) causing people (or units) to not go into the intervention group, unrelated to outcome? This can be used to create a sub-group in the potential comparison group who are more likely to be like the people (units) in the intervention group because the only reason they are not in the intervention group is because of the characteristic. For instance they may live a long way away from where the intervention is going to be held - so the reason they have not chosen to go into the intervention group is not because of something like motivation (which could be related to the outcome) but simply because of travel costs (which it is believed will be unrelated to the outcome).	Instrumental variable	Compare the outcomes of the intervention group with those of a sub-group of the potential comparison group - only those who have the selection characteristic (e.g. they live too far away) and assuming that because they are not in the intervention group just for this reason (i.e. they live too far way), they will be similar to the intervention group in all other important characteristics.
Can describe the intervention group and potential comparison group members accurately so that predictions can be made as to the likely outcome for someone (or some unit) in the absence of receiving the intervention? This prediced outcome (with no intervention) is then compared to the actual outcome which occurred for the intervention group.	Propensity matching	Take the potential comparison group (who have not received any intervention) and using statistical procedures, attempt to predict their outcomes from the characteristics of the group (e.g. age, gender, education, ethnicity, disability). Develop a mathematical formula which predicts the outcome for people (or units) with particular sets of characteristics. For each of the members of the intervention group use the formula to predict what their outcomes would have been likely to be WITHOUT THE INTERVENTION. Compare their actual results on the outcomes (after they have received the intervention) with their predicted results (which would have occurred if they had NOT received the intervention)
Can construct a companison group by locating other individuals (or units) which are exactly 'matched' with members of the intervention group on key characteristics (variables)?	Case matching	For each member of the intervention group, locate individuals who have similiar characteristics to the members of the intervention group apart from not receving the intervention. Compare the outcomes for the intervention group members with their 'matches' in the comparision group.

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D136-0.
Comments for improving this decision-making table welcome - please post at the end of the article referred to immediately above.

doview.com model

Table 2: Selecting techniques to improve constructed matched comparison group designs

Checklists for the seven impact/outcome evaluation designs

Below is a set of checklists, one for each of the seven impact/outcome evaluation designs. They can help the user assess the appropriateness, feasibility and affordability of each type of design. These can be used when designing an impact/outcome evaluation and also when peer reviewing an impact/outcome evaluation which has been set up with a particular type of design.

True experiment design checklist

Appropriateness checklist

☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

☐ Ethical to withhold the intervention from the control group?

☐ Possible to measure the outcome with sufficient accuracy?

Feasibility checklist

☐ Sufficient pool of potential participants (units)? The pool needs to be large enough to get sufficient participants (units) within the timeframe of the evaluation.

☐ Low drop out rates. A sufficient proportion of the intervention and control group must continue for the length of the experiment.

☐ Final sample size large enough? The final sample size needs to be large enough for statistical analysis.

☐ No compensatory intervention for the control group? The control group cannot receive any type of compensatory intervention during the course of the experiment just because they, or outside stakeholders, may want to compensate them for not receiving the intervention.

☐ Appropriate intervention delivered? Need to ensure the appropriate intervention is delivered by either having sufficient control over the quality or quantity of the intervention (intervention fidelity assessment) for the evaluation to be a true test of the intervention.

☐ Sufficient standard of data collection? Needs to be sufficient control over standards of data collection for accurate measurement of outcomes.

☐ Sufficient process description? Need to describe the course and context of the program so that it will be possible to interpret negative evaluation findings (e.g. they may be caused by something such as the intervention was being delivered in an appropriate manner). (Intervention fidelity assessment).

Affordability checklist

☐ Evaluation able to be funded right through to completion? There need to be enough resources (for managing and administering the evaluation, stakeholder engagement, participant selection, interventon process monitoring, data collection, data analysis, trouble shooting, evaluation write up, reporting, and other dissemination) to continue the evaluation right through to its conclusion.

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D138-0.

doview.com model

Checklist 1: True experiment design checklist

Regression discontinuity design checklist

Appropriateness checklist

- ☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

- ☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

- ☐ Ethical to withhold the intervention from the control group? This may be more ethical in a regression discontinuity design because the intervention can be given to those most needing it.

- ☐ Possible to measure a sufficient range of values on selection variable (e.g. outcome)? It needs to be possible to measure a sufficient range of values on the selection variable which is used to decide who will, and who will not, receive the intervention.

Feasibility checklist

- ☐ Sufficient pool of potential participants (units)? The pool needs to be large enough to get sufficient participants (units) within the timeframe of the evaluation.

- ☐ Low-drop out rates. A sufficient proportion of the intervention and control group must continue for the length of the experiment.

- ☐ Final sample size large enough? The final sample size needs to be large enough for statistical analysis.

- ☐ No compensatory intervention for the control group? The control group cannot receive any type of compensatory intervention during the course of the experiment just because they, or outside stakeholders, may want to compensate them for not receiving the intervention.

- ☐ Appropriate intervention delivered? Need to ensure the appropriate intervention is delivered by either having sufficient control over the quality or quantity of the intervention (intervention fidelity assessment) for the evaluation to be a true test of the intervention.

- ☐ Sufficient standard of data collection? Needs to be sufficient control over standards of data collection for accurate measurement of outcomes.

- ☐ Sufficient process description? Need to describe the course and context of the program so that it will be possible to interpret negative evaluation findings (e.g. they may be caused by something such as the intervention not being delivered in an appropriate manner). (Intervention fidelity assessment).

Affordability checklist

- ☐ Evaluation able to be funded right through to completion? Are need to be enough resources (for managing and administering the evaluation, stakeholder engagement, participant selection, interventon process monitoring, data collection, data analysis, trouble shooting, evaluation write up, reporting, and other dissemination) to continue the evaluation right through to its conclusion?

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D139-0.

doview.com model

Checklist 2: Regression discontinuity design checklist

Time series analysis design checklist

Appropriateness checklist

- ☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

- ☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

- ☐ The intervention being applied to everyone? Time series analysis can be used when an intervention is being applied to a whole population.

Feasibility checklist

- ☐ Collection of sufficient number of accurate observations? Is it going to be feasible to collect a sufficiently large number (e.g. 30 or so) measurements of outcomes over time?
- ☐ Accurate measurement of timing of intervention. Is there going to be sufficiently accurate measurement of the timing of the introduction of the intervention so that its timing can be examined in relationship to the observed results in the time series?
- ☐ Appropriate intervention delivered? Need to ensure the appropriate intervention is delivered by either having sufficient control over the quality or quantity of the intervention (or intervention fidelity assessment) for the evaluation to be a true test of the intervention.
- ☐ Sufficient standard of data collection? Needs to be sufficient control over standards of data collection for accurate measurement of outcomes.
- ☐ Sufficient process description? Need to describe the course and context of the program so that it will be possible to interpret negative evaluation findings (e.g. they may be caused by something that the intervention was not delivered in an appropriate manner). (Intervention fidelity assessment).

Affordability checklist

- ☐ Can measurement be continued long enough. Are the sufficient resources for measurement to be continued long enough to get a full time series?
- ☐ Evaluation able to be funded right through to completion? Are there enough resources (for managing and administering the evaluation, stakeholder engagement, participant selection, interventon process monitoring, data collection, data analysis, trouble shooting, evaluation write up, reporting, and other dissemination) to continue the evaluation right through to its conclusion?

doview.com model

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D140-0.

Checklist 3: Time series analysis design checklist

Constructed matched comparison group design checklist

Appropriateness checklist

☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

☐ Ethical to encourage the intervention to be withheld from the comparison group?

☐ Intervention group sufficiently similar to comparison group? The information collected on the members of the compaison group needs to be sufficient to make sure that they are fully matched to members of the control group. This can be a problem where it is suspected that there are subtle biases operating in individuals entering the intervention group which are not able to be measured by the data collected on the comparison group (e.g. the level of motivation of those entering the intervention group).

☐ Same conditions operating in different time periods? If the information describing those in the comparison group is coming from a different time period, then there cannot be anything significantly different about that time period versus the time period in which the intervention group information is being collected.

Feasibility checklist

☐ Sufficient pool of potential participants. Will there be a sufficient pool of participants available for selecting the intervention and comparison groups within the timeframe of the evaluation?

☐ No compensatory intervention for the control group? The control group cannot receive any type of compensatory intervention during the course of the experiment just because they, or outside stakeholders, may want to compensate them for not receiving the intervention.

☐ Low drop out rates. A sufficient proportion of the intervention or control group must continue for the length of the evaluation.

☐ Final sample size large enough? The final sample size needs to be large enough for statistical analysis.

☐ Appropriate intervention delivered? Need to ensure the appropriate intervention is delivered by having sufficient control over the quality or quantity of the intervention 9or intervention fidelity assessment) for the evaluation to be a true test of the intervention.

☐ Sufficient standard of data collection? Needs to be sufficient control over standards of data collection for accurate measurement of outcomes.

☐ Sufficient process description? Need to describe the course and context of the program so that it will be possible to interpret negative evaluation findings (e.g. they may be caused by something that the intervention was not delivered in an appropriate manner). (Intervention fidelity assessment).

Affordability checklist

☐ Evaluation able to be funded right through to completion? Needs to be enough resources (for managing and administering the evaluation, stakeholder engagement, participant selection, interventon process monitoring, data collection, data analysis, trouble shooting, evaluation write up, reporting, and other dissemination) to continue the evaluation right through to its conclusion.

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D141-0.

doview.com model

Checklist 4: Constructed matched comparison group design checklist

Exhaustive causal identification and elimination design

Appropriateness checklist

☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

☐ Clear defintion of success? It must be clear when their has been success on the outcome measure.

☐ Sufficiently rich information about the intervention being delivered and causal mechanisms? Is it likely to be possible to collect sufficiently rich information about the possible mechanisms through which the alternative potential causes could have caused the outcomes to occur. (I.e. the situation needs to be one where stakeholders,or the evaluator, wil be able to have available sufficient information to make reasonably accurate judgements about alternative causal mechanisms).

Feasibility checklist

☐ Sufficient pool of participants(units) to study? Will there be a sufficient pool of participants (units) available to study for whom outcomes have been positive (or negative if this group is also being studied) within the timeframe of the evaluation?

☐ Accurate measurement of outcomes? Will it be feasible to accurately measure the outcomes for those who have received the intervention? (In technical terms when first taken this measurement is called a 'not-necessarily demonstrably attributable indicator'). It is needed in this design so that one can examine cases in which outcomes have occured to eliminate the possible alternative causal explanations. Cases where the outcome has not been achieved may also be studied.

☐ Sufficient standard of data collection? Needs to be sufficient control over standards of data collection for accurate measurement of outcomes.

☐ Possible to identify alternative causes? I must be likely that it will be possible to exhaustively identify all possible alternative causes.

☐ Elimination of alternative causes? Is it likely that it will be possible to have available enough information to clearly eliminate alternative causes as having been the cause of outcomes to be achieved.

☐ Sufficient process description? Need to describe the course and context of the program so that it will be possible to interpret negative evaluation findings (e.g. they may be caused by something that the intervention was not delivered in an appropriate manner). (Intervention fidelity assessment).

Affordability checklist

☐ Evaluation able to be funded right through to its completion? Are there enough resources (for managing and administering the evaluation, stakeholder engagement, participant selection, interventon process monitoring, data collection, data analysis, trouble shooting, evaluation write up and reporting, other dissemination) to continue the evaluation right through to its conclusion?

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D142-0.

doview.com model

Checklist 5: Exhaustive causal identification and elimination design

Expert judgment design checklist

Appropriateness checklist

☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

☐ The right evaluation question being asked and answered? Answering the wrong evaluation question can be a waste of effort.

☐ Topic one in which there are recognized experts? Is the topic one in which there are recognized experts who could be contracted to undertake an evaluation based on their expert knowledge.

☐ Expert(s) able to make judgments on program? Is the topic one on which experts are likely to be able to make a judgment as to whether, or not, the program has improved outcomes. This will require that causal mechanisms are likely to be sufficiently clear to experts and that there are not a large number of credible alternatives as to what has affected outcomes.

☐ Key stakeholders going to accept this as sufficiently robust design? Key stakeholders need to accept that this evaluation method will provide them with sufficiently robust findings in regard to the impact of the program for the purposes for which they want to use such findings and the level of resources they have to spend on the evaluation.

Feasibility checklist

☐ Expert(s) sufficiently free from bias? Will it be possible to locate expert(s) who are sufficiently free from bias (either in favor of, or opposition to, the program) such that they can provide an credible judgment on the outcome of the program?

☐ Expert(s) judgments accepted by stakeholders? Will the judgment of the expert(s) be accepted by a sufficient number of key stakeholders as providing a credible judgment on whether or not the program improved outcomes.

Affordability checklist

☐ Evaluation able to be funded right through to its completion? Are there enough resources (for managing and administering the evaluation, stakeholder engagement, employing the expert(s), paying to collect whatever information the expert(s) will need to make their judgement, trouble shooting, evaluation write up and reporting, other dissemination) to continue the evaluation right through to its conclusion?

doview.com model

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D143-0.

Checklist 6: Expert judgment design checklist

Key informant judgment design checklist

Appropriateness checklist

☐ The right evaluation question being asked and answered? Based on what is already known in the sector and the priority evaluation questions to build sector knowledge.

☐ The best project on which to answer the evaluation question? There may be a better project in the sector (more typical, more innovative, more controllable) on which to answer the evaluation question and evaluation resources might be pooled from various projects to answer it properly there.

☐ The right evaluation question being asked and answered? Answering the wrong evaluation question can be a waste of effort.

☐ Key informants able to make judgment on program? Is the topic one on which key informants are likely to be able to make a judgment as to whether, or not, the program has improved outcomes. This will require that causal mechanisms are likely to be sufficiently clear to stakeholders and that there are not a large number of credible alternatives as to what has affected outcomes.

☐ Key stakeholders going to accept this as sufficiently robust design? Key stakeholders need to accept that this evaluation method will provide them with sufficiently robust findings in regard to the impact of the program for the purposes for which they want to use such findings and the level of resources they have to spend on the evaluation.

Feasibility checklist

☐ Key informants willing to be involved in making judgements? Will there be a sufficiently wide range of key informats willing to make judgments on whether the program improved outcomes.

☐ Key informants sufficiently free from bias. Will key informant judgments be sufficiently free from bias (either in favor of, or opposition to, the program) such that they can provide an credible judgment on the outcome of the program?

☐ Key informant judgments accepted by stakeholders? Will the judgments key informants be accepted by a sufficient number of key stakeholders as providing a credible judgment on whether or not the program improved outcomes.

Affordability checklist

☐ Evaluation able to be funded right through to its completion? Are there enough resources (for managing and administering the evaluation, stakeholder engagement, interviewing or otherwise involving the key informants, trouble shooting, evaluation write up and reporting, other dissemination) to continue the evaluation right through to its conclusion?

doview.com model

Duignan, P. (2009) Selecting impact/outcome evaluation designs - a decsion-making table and checklist approach. Outcomes Theory Knowledge Base Article # 256 (OutcomesTheory.org) D144-0.

Checklist 7: Key informant judgment design checklist

Conclusion

Two decision-making tables and a set of seven checklists have been set out to provide a framework for selecting impact/outcome evaluation designs and for selecting techniques to improve constructed matched comparison group designs. They can be used in impact/outcome evaluation planning and in peer reviewing evaluations which have been planned by others.

Please comment on this article

This article is based on the developing area of [outcomes theory](#) which is still in a relatively early stage of development. Please critique any of the arguments laid out in this article so that they can be improved through critical examination and reflection.

Acknowledgments

This decision-making table approach was developed by the author as a result of his involvement as an expert evaluation adviser in the [Youth Employment Network](#) (YEN) (a partnership between the International Labor Organization, the World Bank and the United Nations) Evaluation Clinic held in Damascus, Syria on 19-20 July 2009. In particular he benefited from discussions with other expert advisers on techniques to improve constructed matched comparison group designs: Rita Almeida (Economist, World Bank), David Newhouse (Economist, World Bank), Mattias Lundberg (Senior Economist, World Bank), Alexandre Kolev (ILO/International Training Centre (ITC)); YEN staff, Suzanna Puerto (Technical Officer), Marcus Pilgrim (Manager) and Drew Gardiner (Technical Officer); Nader Kabbani (Director of Research, Syria Trust for Development); and as a result of working with clinic participants on possible impact/outcome evaluation designs with participants.

Citing this article

Duignan, P. (2009). *Selecting impact/outcome evaluation designs: a decision-making table and checklist approach*. Outcomes Theory Knowledge Base Article No. 256. (http://knol.google.com/k/paul-duignan-phd/selecting-impactoutcome-evaluation/2m7zd68aaz774/115).

[If you are reading this in a PDF or printed copy, the web page version may have been updated].

[Outcomes Theory Knowledge Base Article # 256 http://www.tinyurl.com/ot256]

Comments

Write New Comment ▼

Selecting impact/outcome evaluation designs: a decision-making table and checklist approach – a knol by Paul Duignan, PhD

15/08/09 7:32 AM

[Ausevaluator](#)
[Invite as author](#)

Untitled

Hi Paul, I really like the idea of looking hard at the intervention and the evaluation situation and then selecting an approach to match. I have 2 comments about the current version. For the second column (about 'contamination') I don't understand why the language is different to the other columns. I think it would be better to stick to the label 'May be appropriate' rather than 'No control group'.

I've also been thinking about the list of designs and what I don't see is consideration of other analytical strategies like looking for a dose-response, checking temporality, and arguing by analogy.

Finally, I wonder if we can be careful when we're discussing 'ruling out alternative explanations' that we don't exclude the possibility that 2 different causes both contributed.

Patricia Rogers

Last edited Aug 13, 2009 3:09 PM
[DeleteBlock this userReport abusive comment](#)
0 [View/post replies (1) to this comment ▼](#)