



TRANSLATING RESEARCH
INTO ACTION CENTER

AMERICAN UNIVERSITY • WASHINGTON, DC

Measuring the Societal Impact of ART

**NSF Accelerating Research Translation (ART) Program
Award #2331399**

Prof. Susanna P. Campbell
Provost Associate Professor, School of International Service
American University

September 27, 2024

TRAC is funded, in part, through the National Science Foundation Accelerating Research Translation (ART) Program Award #2331399. The opinions, findings, conclusions, or recommendations expressed here within are attributed to the author(s) and do not necessarily reflect the views of the National Science Foundation

Overview

- Why Measurement Matters
- Components of Monitoring and Evaluation
- Challenge to Measuring ART's Societal Impact
- Basic Research Design Considerations
- Options for Monitoring and Evaluating ART

Why Measuring Societal Impact Matters



Enables learning and adaptation, improving impact



Captures engagement necessary for impact, enabling IHEs to reward engagement



Supports justification of return on investment

Components of Monitoring

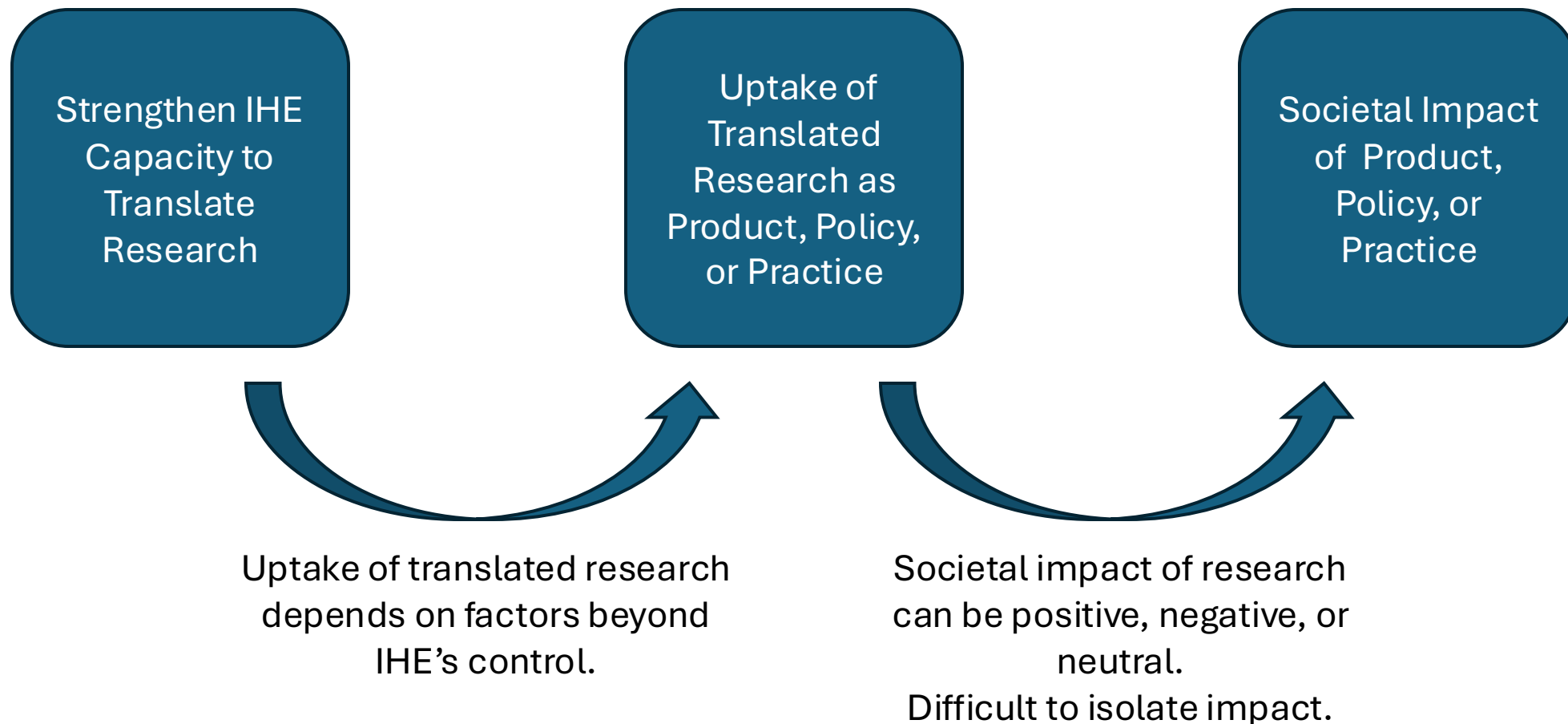
- Monitoring = gathering data on:
 - Inputs:
 - What the project does (e.g., implements T&P reforms)
 - How it does this (e.g., engages a broad group of faculty with the aim that reforms are broadly accepted)
 - Intermediary outcomes:
 - Progress toward the project outcomes (e.g., adoption of the T&P guidelines by a new department)
 - Outcomes:
 - Change that aligns with the project objectives (e.g., use of revised T&P guidelines in all tenure and promotion decisions).
 - Impact:
 - Longer-term change aligned with project goals and broader aims (e.g., change in culture of T&P within the university and discipline)
- Why monitor?
 - If you do not know what you did, or how you did it, you will not be able to take credit for doing it, learn from it, or assess its impact.

Types of Evaluation

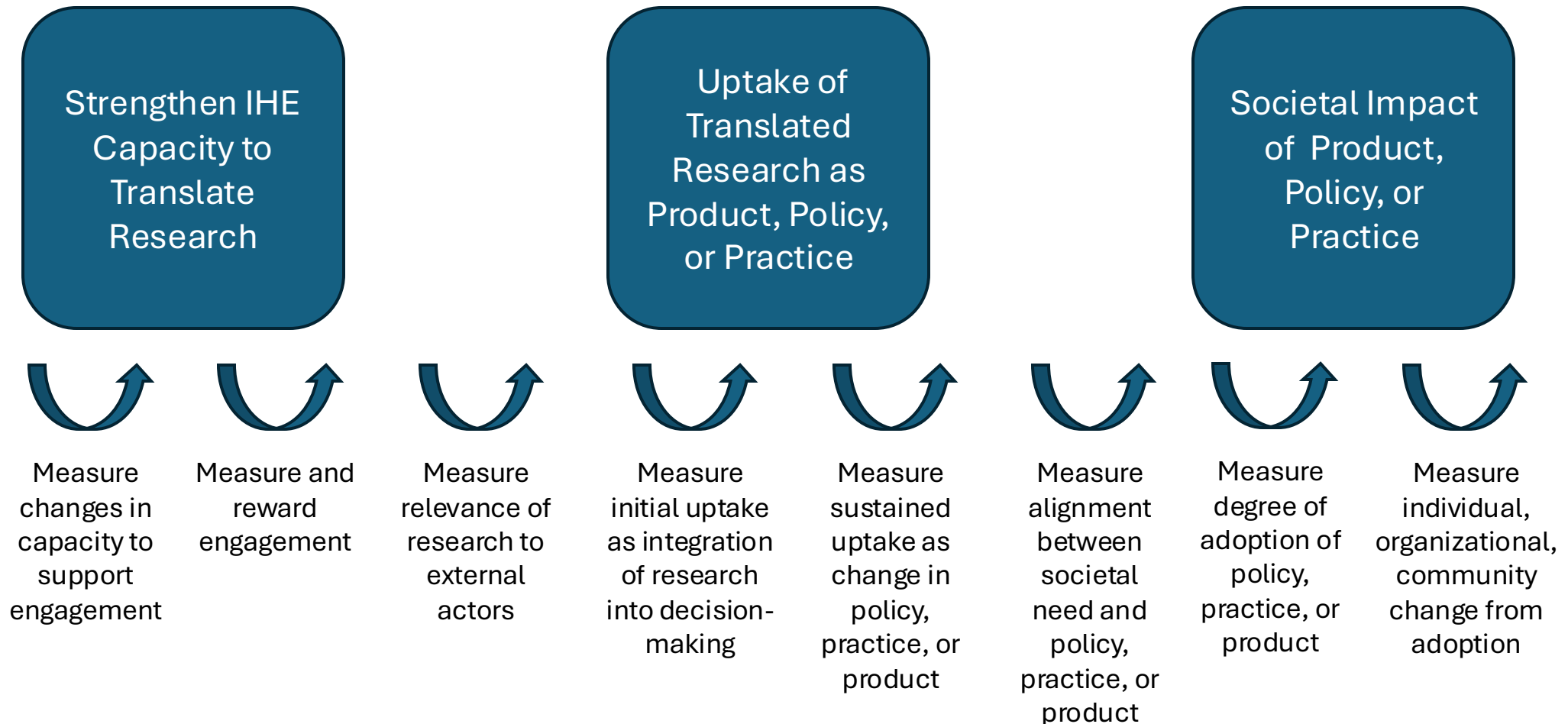
- Assessment of intermediary outcomes with the aim of improving impact:
 - Formative evaluation – assess feasibility of project and suggest changes
 - Developmental evaluation – provide feedback during implementation
 - Process evaluation – assess how project was implemented
- Assessment of impact:
 - Randomized Controlled Trial – randomize assignment of treatment
 - Quasi-experimental design – “as if” random assignment
 - Medium-N theory testing – necessary/sufficient conditions for outcomes
 - Case studies – process tracing + comparative cases + participatory methods

Resource: <https://www.3ieimpact.org/pir-methods>

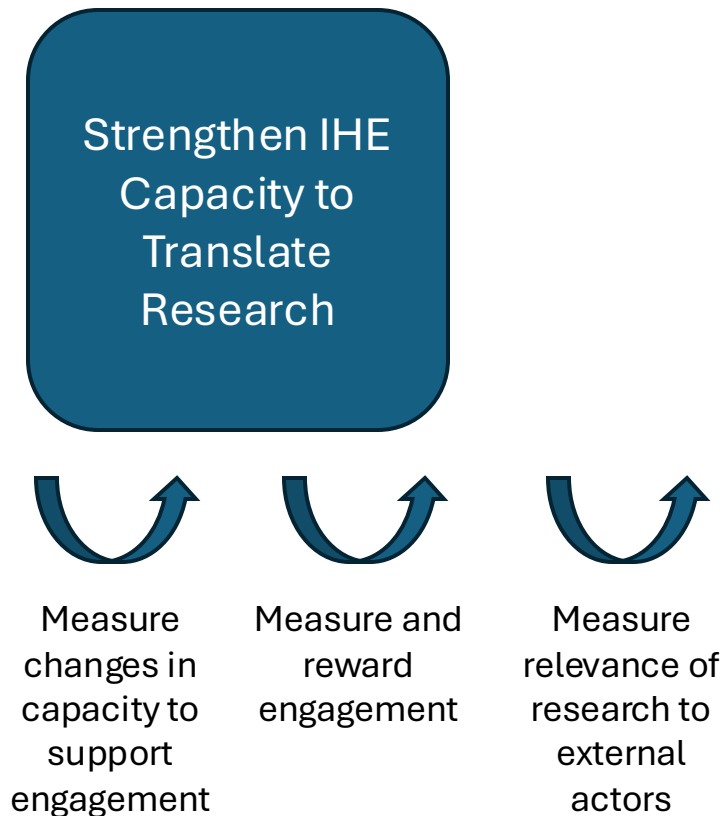
Challenge of Measuring ART's Societal Impact



Solution to Measuring ART's Societal Impact



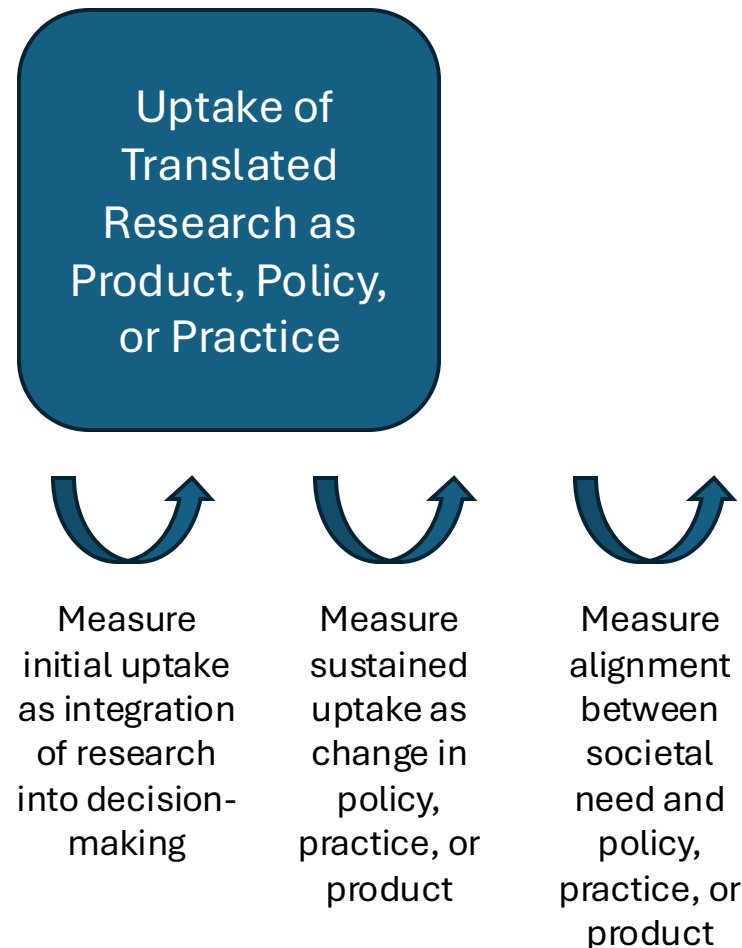
Measuring IHE Capacity to Translate Research



- *Inputs:*
 - Changes in IHE capacity to support external engagement related to research.
- *Intermediary outcome:*
 - Engagement with policymakers, practitioners, communities, private sector actors.
- *Outcomes:*
 - High relevance of research outputs to external actors.

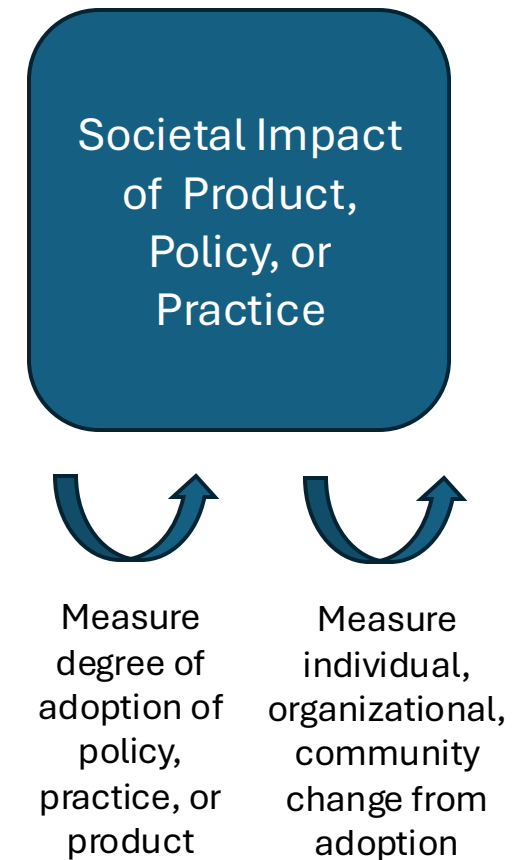
Measuring Uptake of Translated Research

- *Inputs:*
 - Research translated into format accessible to external actor.
- *Intermediary outcomes:*
 - Findings used in decision-making process.
- *Outcomes:*
 - Change in policy, practice, or product related to findings.
 - Changes in policy, practice, or product relevant to (equitable) societal need.



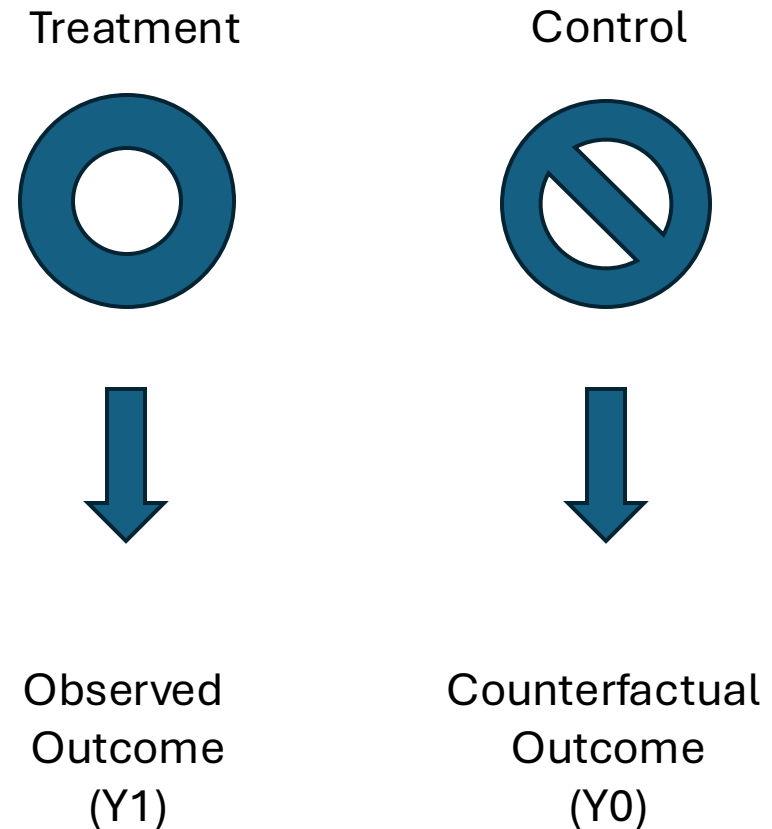
Measuring Societal Impact of Research Uptake

- *Inputs:*
 - Implementation of policy, practice, or product (related to findings).
 - Integration of strategy to improve broader societal impact.
- *Outcomes:*
 - Changes at the individual, organization, community, etc... levels related to implementation.
- *Impact:*
 - Longer-term changes at the individual, organization, community, etc... levels related to implementation.



Fundamental Problem of Causal Inference

- We cannot simultaneously observe a person in its treated and untreated states
 - i.e. we cannot observe the *counterfactual*.
- We need to *infer causality* using observational or experimental research designs.



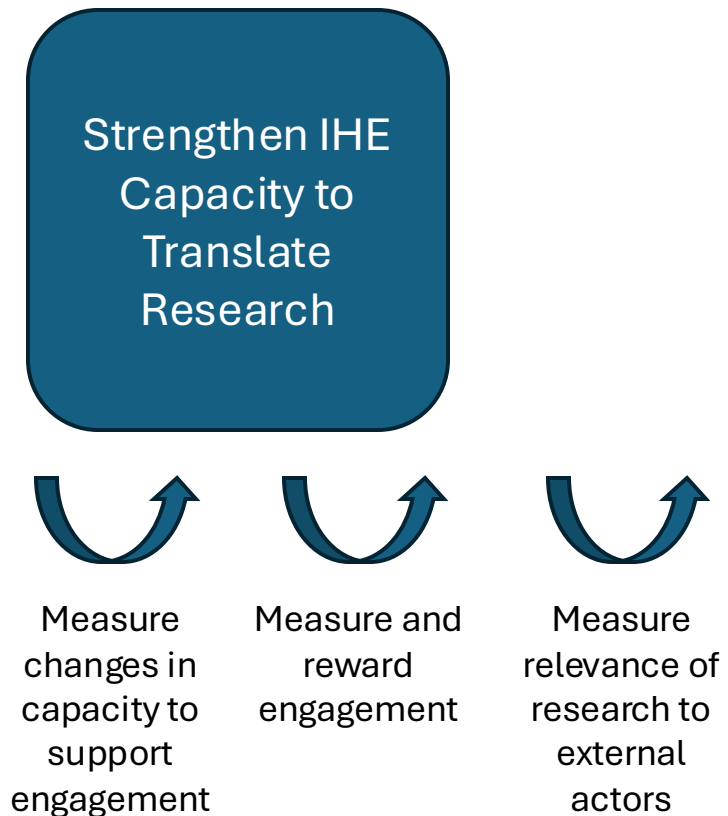
Experimental vs. Observational Designs

- *Experimental research* – can manipulate the allocation of treatment through random assignment.
 - Random assignment of treatment eliminates the selection bias.
 - Average Treatment Effect (ATE) = $E[Y_1] - E[Y_0]$
- *Observational research* – cannot “manipulate” the allocation of the treatment. Need to find a *comparison group* that has not received the treatment and can be used to infer causality.
 - *Small-N case design*: compare cases that allow you to trace and validate the causal chain and rule out alternatives.
 - *Medium-N design*: compare all cases in your population to identify necessary and sufficient conditions.
 - *Quasi-experimental designs*: match participants and non-participants using “as if” random strategy - matching, closeness to eligibility, etc.

Probabilistic vs. Deterministic Causality

- Probabilistic causality
 - **Observational** studies: X on average causes Y (in a large population), given that the following criteria are met:
 - *Association*: changes X are accompanied by systematic changes Y.
 - *Direction of influence*: from cause to effect, i.e. a change in X brings about a change in Y.
 - *Non-spuriousness*: the two associated variables X and Y have no common extraneous cause Z.
 - **Experimental** studies: X on average causes Y (for the treatment population)
 - Measure the average impact of having access to the program (*intention-to-treat*).
- Deterministic causality
 - The relationship between the variables is known for certain within the full population falling within specific scope conditions.
 - *Necessary cause*: X must ALWAYS precede the effect Y.
 - *X is a cause of Y, i.e. without X, Y would not have occurred.*
 - *Sufficient cause*: X is SUFFICIENT to produce Y.
 - *X is enough to cause Y.*

Monitoring and Evaluation of IHE Capacity Strengthening



- *Monitor (What and How?):*
 - Training – Initial training, follow-up, and institutionalization
 - Engagement – efforts to engage externally to support research translation
 - Capacity strengthening – build communications, tech transfer, etc...
 - Institutional reforms – T&P, etc...
- *Evaluate (To What end?):*
 - Changes in knowledge
 - Changes in behavior
 - Changes in application of promotion standards
 - Changes in research related to policy/private interests
- *Some Evaluation Design Options:*
 - Case studies of training or specific capacity building (embed survey)
 - Causal process tracing – testing theory of capacity building
 - Observational survey (changes in reported capacity)
 - Developmental evaluation – support learning

Monitoring and Evaluation of Uptake of Translated Research

- *Monitor (What and How?):*

- How is the research translated (methods)?
- Who uses research?
- How do they use it?

- *Evaluate (To What end?):*

- Inclusion of findings in decision-making process, practice discussions, or product development.
- Change in policy, practice, or product related to findings.
- Relevance of policy, practice, or product to societal need.



Measure initial uptake as integration of research into decision-making



Measure sustained uptake as change in policy, practice, or product



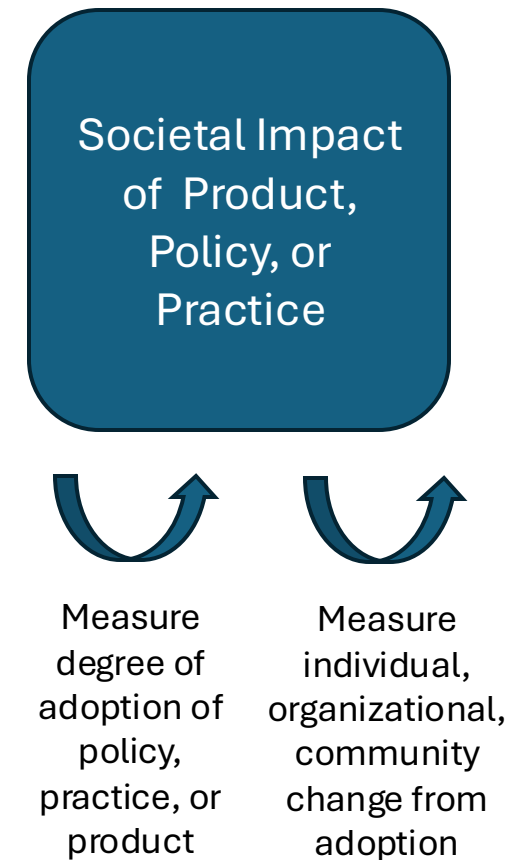
Measure alignment between societal need and policy, practice, or product

- Some Evaluation Design Options:

- RCT:
 - Randomize exposure to findings
 - Randomize method of exposure to findings
- Quasi-Ex:
 - RDD – Cut-off around training eligibility
- Large-N Observational:
 - Survey of discourse changes
 - Text analysis of public comms or code cables
- Medium/Small-N:
 - Case studies with process tracing related to policy/practice/product changes.
 - Content analysis of documents
- Learning:
 - Development Evaluation

Monitoring and Evaluation of Impact of Research Uptake

- *Monitor (What and How?)*
 - Degree to which research is integrated into product, policy, or practice.
- *Evaluate (To What end?):*
 - Variation in implementation/adoption of policy, practice, or product.
 - Related changes in individual, organizations, or communities related.
- **Some Evaluation Design Options :**
 - RCT – Randomly assign practice/product
 - Quasi-Ex – PSM + Diff-in-Diff for treatment/control
 - Medium-N – Compare case population of recipients, or recipients/non-recipients in scope.
 - Small-N – Test theory of change via process-tracing and comparative case.



Thank you!

trac@american.edu

Updates on 2024-2025 Webinar Series Forthcoming.

Next meeting of ART Measurement Community of Practice

November 15, 2024

12-1pm ET

Appendix

Comparison	Large-n Observational	Comparative case study (between-case)	Mechanism (within-case)	Experimental	Interpretivist/ Constructivist
Basis of inference	Regularity Probabilistic	Regularity Deterministic	Laws about how the parts of the mechanism are connected; solving pairing problem	Manipulation/Directionality (agency) Randomization (independence from other factors) Non-interference/ SUTVA (no spillover)	Reflexivity/ Co-constitution
Assumptions underlying inference	Use controls to create an "as-if" counterfactual; Constant conjunction	Use case selection and alternative explanations to create "as-if" counterfactual	Validation of parts of mechanism in case, pairing up X and Y.	Pre-emption (manipulation did not "pre-empt" another cause)	Identification of intersubjective understanding; and co-constitution between structure & Agency
Types of Research Questions	Effects-of-causes What is the (average) effect of variation in X on variation in Y?	Causes-of-effects What explains THIS variation in Y, exhibited in this sub-set of cases?	Explains HOW X connects to Y	Effects-of-causes What is the effect of X treatment on Y outcome?	<i>Constructivist</i> : How do X1 and X2 co-constitute Y context/outcome? <i>Interpretivist</i> : What is THIS process/culture/relationship? <i>Critical</i> : How does X critical theory explain the ideology/belief set/power relationship in THIS case?
Identification of Theory	Deductive/ Inductive	Deductive/ Inductive	Deductive/ Inductive	Deductive	Inductive/ Abductive (iterative)
Design Options	Bivariate, multivariate, multiple regression, etc... Regression Discontinuity Design; Statistical Matching; Difference-in-Difference	QCA; Typology; Most-similar; Most-different; Diverse; Longitudinal Single cases (pathway/typical; influential; extreme; index; deviant) Comparative Historical Analysis	Causal process observations/ process-tracing	Randomized Controlled Trial	Within case & Cross-case analysis
Internal Validity/ External Validity	- Weak internal validity (association) - External validity relies on capturing population	- Weak internal validity (association) - External validity relies on alternative explanations & scope	- Strong external validity relies on consistent causal mechanism - External validity relies on claims of analytic generalization to theories	- Internal validity strong via manipulation, randomization, and non-interference - External validity weak, relies on analytic generalization to theories/cases	- Internal validity – relies on ethnography and thick description - External validity (if desired) relies on alternative explanations and scope conditions
Data and Analysis	Existing datasets Surveys Regression Analysis	Ethnography Semi-structured interviews Archival Structured, Focused Comparison QCA	Ethnography Semi-structured interviews Archival Process tracing Content analysis	Surveys Existing datasets (covariates) Regression (ATE)	Ethnography Semi-structured interviews Archival Content analysis Discourse analysis