Integrating Potential Outcomes and Causal Mechanisms to Guide Multi-Method Research

Marco Alcocer[1]
Harvard University

Can evidence seeking to estimate causal effects and evidence attempting to uncover causal mechanisms be integrated in multi-method research? And if so, how? While some unified frameworks have been presented to guide the integration of causal effects and causal mechanisms in a single study, these use mono-method (quantitative) frameworks and incorporate causal mechanisms that are defined as random intervening or mediating variables (Imai et al. 2011; Glynn and Quinn 2011; Humphreys and Jacobs 2015). Yet, most theory-based and qualitative scholars argue that causal mechanisms are not random variables but static, invariant factors that should be examined through within-unit qualitative methods (e.g., Hedström and Ylikoski 2010; Beach and Pedersen 2019; Goertz and Mahoney 2013; Waldner 2016). Despite this, the literature has yet to provide formal frameworks to guide multi-method research that incorporates qualitative and quantitative methods to investigate causal effects and causal mechanisms. Instead, multi-method research tends to draw on other frameworks, such as the potential outcomes (PO) and causal graphs frameworks, to informally discuss how qualitative methods can be combined with quantitative results (e.g., Psillos 2004; Paluck 2010; Seawright 2016; Goertz 2017).

This project contributes to this literature by presenting a novel unified formal framework to conduct multi-method research. In this short article, I draw on the PO framework (Neyman 1923; Rubin 1974) and extend it to incorporate invariant causal mechanisms. This framework clarifies the role of quantitative and qualitative methods when investigating causal claims in a multi-method study.

While I only present the setup of the framework here, in the larger project I use the framework to derive the role of quantitative and qualitative methods in multi-method research for some of the most popular research designs for applied researchers: including simple randomized experiments, instrumental variables, difference-in-differences, and regression discontinuity. I also discuss key implications of the framework, including the meaning of "counterfactuals" for causal mechanisms, mechanistic heterogeneity, case selection, and generalization of causal mechanisms. Potential Outcomes and Causal Mechanisms

To begin, let us discuss the PO framework by drawing from Morgan and Winship (2015). For a binary case, each unit i has two potential outcome random variables, Y_i^1 in the treatment state and Y_i^0 in the control state. The individual causal effect for unit i is therefore $\theta_i = Y_i^1 - Y_i^0$. For each causal state, a treatment or exposure variable D_i exists, where $D_i = 1$ for units exposed to the treatment state and $D_i = 0$ for units exposed to the control state. If we assume that some mechanism exists that leads the treatment variable D_i to cause a change in Y^1, then we have identified where causal mechanisms fit into the potential outcome framework. Accordingly, I expand the PO framework to accommodate causal mechanisms.

For unit i, $D_i = 1$ causes a change in Y_i^1 through $M_i(D_i = 1)$, which is a non-empty set of mechanisms $M_i(D_i = 1) = \{m_1, m_2, m_3, \ldots, m_n\}_i$. The set, $M_i(D_i = 1)$, has at least one mechanism, m_n, and if there are more than one, the mechanisms need not be mutually exclusive. In other words, $D_i = 1$ can cause Y_i^1 through more than one mechanism, and maybe even through *a combination* of these mechanisms. For example, $M_i(D_i = 1)$ may cause Y_i^1 through m_1, or through $m_1 \wedge m_2$, or through $(m_1 \wedge m_2) \vee m_3$.[2] Importantly, $D_i = 0$ does not have any mechanisms since it is not causing anything, and therefore $M_i(D_i = 0) = \varnothing$. This implies that mechanisms are only realized when $D_i = 1$.

However, the fundamental problem of causal inference is that we cannot observe both potential outcomes. That is, for each unit, only one of the potential outcomes is realized, so that the observed outcome variable is

$$Y_i = \begin{array}{l} Y_i^1 \text{ if } M_i(D_i = 1) \\ Y_i^0 \text{ if } M_i(D_i = 0) \end{array} \quad (1)$$

Written differently,

$$Y_i^0 = M_i(D_i)Y_i^1 + (1 - M_i(D_i))Y_i \quad (2)$$

1 I thank Jason Seawright for his guidance, mentorship, and feedback throughout the process of this project. I thank Stephan Haggard for his feedback on early drafts of the project. I thank the participants of the 2023 Emerging Methodologists Workshop for their insights, Daniel Pemstein for his comments, and Diana Kapiszewski, Hillel David Soifer, and Qualitative and Multi-Method Research section of APSA for supporting the project.

2 Here $\wedge$ means "and", while $\vee$ means "or".

Notice that because the causal states Y_i^1 and Y_i^0 are unrealized, no mechanisms exist in these states. It is only when the treatment variable D_i is realized that mechanisms are also realized, but only for units where $D_i = 1$. In other words, the extended PO framework can explain how the realized potential outcome, Y_i, and the treatment variable, D_i, are random variables while mechanisms are static and invariable within each unit i (M_i is only realized when $D_i = 1$ and not when $D_i = 0$).

Moreover, if we believe that causal effects are probabilistic (not deterministic) or that outcomes are not monocausal, then for realized cases of Y_i (where $Y_i = 1$ denotes the outcome is realized and $Y_i = 0$ means the outcome is not realized), we will likely observe four types of general observations in our realized data: $Y_i(M_i(D_i = 1)) = 1$, $Y_i(M_i(D_i = 1)) = 0$, $Y_i(M_i(D_i = 0)) = 1$, or $Y_i(M_i(D_i = 0)) = 0$. Given the framework presented here, this means that mechanisms exist for both $Y_i(M_i(D_i = 1)) = 1$ *and* $Y_i(M_i(D_i = 1)) = 0$. In the case that $Y_i(M_i(D_i = 1)) = 1$, the causal mechanisms realized by $D_i = 1$ should link D_i to Y_i. However, in the case that $Y_i(M_i(D_i = 1)) = 0$, we should observe some factor disrupting the causal mechanisms realized by $D_i = 1$ that should have caused an effect on Y_i but does not.

An important question that remains is whether the causal mechanism(s) is the same across units, or $M_i(D_i = 1) = M(D_i = 1)$ for all $i = 1, \ldots, n$ (mechanistic homogeneity), whether it varies across subsets of the units (mechanistic homogeneity within subgroups and mechanistic heterogeneity across subgroups), or whether M_i is unit-specific (complete mechanistic heterogeneity). This question is equivalent to asking whether there are constant causal effects ($Y_i^1 - Y_i^0 = \theta$ for all i) or heterogeneous causal effects ($Y_i^1 - Y_i^0 = \theta_i$), except focusing on mechanisms. In the social sciences our theories most often tend to assume mechanistic homogeneity or mechanistic homogeneity within a subset of units—for example, the effect of economic development on democratization varies by level of economic inequality (low, medium and high). We certainly never assume complete mechanistic heterogeneity.

While the PO framework that incorporates causal mechanisms is presented here using a binary treatment condition, $D_i \in \{0, 1\}$, the framework can be extended to non-binary treatment. Further, for simplicity, like the basic PO setup, I also make the stable unit treatment value assumption (SUTVA).

Future Research and Discussion

In the broader project, I take the new framework that incorporates invariant causal mechanisms and use it to derive the role of causal mechanisms when we use quantitative methods to estimate the average causal effects, including simple randomized experiments, linear regression, difference-in-differences, instrumental variables, and regression discontinuity designs. This identifies the role of causal mechanisms and qualitative methods in multi- method research when these quantitative tools are used. I also discuss key implications of the framework in detail, including "counterfactuals" for causal mechanisms, mechanistic heterogeneity, case selection criteria, and the generalizability of causal mechanisms in multi-method research.

In sum, the framework presented here provides not only a theoretical but a practical guideline for conducting multi-method research for causal claims. This framework has the potential to guide more rigorous and robust multi-method research. It also advances the ontological and epistemic underpinnings of multi-method research and contributes to the growing literature on this methodological approach.

References

Beach, Derek and Rasmus Brun Pedersen (2019). *Process-tracing methods: Foundations and guidelines.* University of Michigan Press.

Glynn, Adam N and Kevin M Quinn (2011). Why process matters for causal inference. *Political Analysis 19* (3), 273–286.

Goertz, Gary (2017). *Multimethod research, causal mechanisms, and case studies: An integrated approach.* Princeton University Press.

Goertz, Gary and James Mahoney (2013). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences.* Princeton: Princeton University Press.

Hedström, Peter and Petri Ylikoski (2010). Causal mechanisms in the social sciences. *Annual review of sociology 36*, 49–67.

Humphreys, Macartan and Alan M Jacobs (2015). Mixing methods: A Bayesian approach. *American Political Science Review 109* (4), 653–673.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review 105* (4), 765–789.

Morgan, Stephen L and Christopher Winship (2015). *Counterfactuals and causal inference.*

Cambridge University Press.

Neyman, Jerzy (1923). On the application of probability theory to agricultural experiments. essay on principles. *Ann. Agricultural Sciences*, 1–51.

Paluck, Elizabeth Levy (2010). The promising integration of qualitative methods and field experiments. *The ANNALS of the American Academy of Political and Social Science 628* (1), 59–71.

Psillos, Stathis (2004). A glimpse of the secret connexion: Harmonizing mechanisms with counterfactuals. *Perspectives on science 12* (3), 288–319.

Rubin, Donald B (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66* (5), 688.

Seawright, Jason (2016). *Multi-method social science: Combining qualitative and quanti- tative tools*. Cambridge University Press.

Waldner, David (2016). Invariant causal mechanisms. *Qualitative & Multi-Method Research 14* (1-2), 28–33.

Qualitative and Multi-Method Research | Fall 2023, Volume 21.2 | https://doi.org/10.5281/zenodo.8418917

Measuring Costly Concepts: Validation Samples for Measuring Many-N Cases

Sarah Moore
Northwestern University

Costly concepts are concepts that are expensive or otherwise resource-intensive to obtain measurement for over many cases. Costly concepts are present across the social sciences, though particularly in the subnational study of comparative politics. Subnational democracy, local-level armed group presence, and municipal corruption are all costly concepts for which measurement requires fine-grained data that may be practically impossible to collect for many units where the data are not already available to researchers. In the absence of actual measures of costly concepts, scholars will often substitute measurement by using proxy variables in empirical analyses, which causes non-random measurement error where measurements of the costly concept and proxies are not identical. This non-random measurement error means we risk conducting biased analyses when we cannot overcome the structural challenges that preclude precise measurement of costly concepts.

For example, the quantitative literature on non-state armed actors and violent conflict has overwhelmingly relied on local violence data to measure the presence of armed groups throughout a territory (for more extensive reviews of this literature, see Arjona and Castilla 2022; Vela Barón 2021)for obvious reasons, on violence. Yet, civil war is about much more than violence. We argue that the focus on violence hinders our understanding of the most common type of armed conflict in the world today. In particular, equating civil war and violence leads to (i. However, measuring armed group presence through violence fails as a proxy in ways we would easily expect given existing theory on civil war violence (Arjona 2016; Kalyvas 2000). Alternative measures of armed group presence entail gathering extensive knowledge from local experts through fieldwork (e.g., Arjona 2016; Aponte-González, Hirschel-Burns, and Uribe 2023). However, fieldwork-based approaches to measuring local-level armed group presence are incredibly expensive and thus limited to a reduced number of cases.

How do we know the extent to which a proxy can reliably substitute measures of our costly concept? How do we improve proxies or other measures when the proxy alone is unreliable? In this work, I develop methodological tools to understand the performance of existing proxies for costly concepts and inform more sophisticated measurement strategies based on the direct measurement of a subset of cases where obtainable. Here, I focus on a summary of the former, in which I develop a framework for collecting and analyzing validation samples wherein the accurate measurement of the costly concept is obtained for a set of cases to discern the performance of a proxy over three dimensions: the extent of disagreement, the variation in the disagreement, and the predictive features of the disagreement. I further assess the type of sample required to best estimate proxy performance relative to three potential options: a random sample, a stratified random sample, or a theoretically informative sample.

My overarching argument is that having at least some information about the relative performance of a potential proxy is better than uninformed analysis with said proxy. Collecting validation samples of at least a subset of cases to obtain direct measurements of a costly concept allows researchers to understand the degree to which a proxy and concept of interest converge and provides insight into the circumstances where they do not. To illustrate the proposed methodological framework and discuss the trade-offs of some of the sampling approaches available for these validation samples in the larger paper, I rely on simulated data. I use the concept of armed group