Jacobs, Alan M. and Tim Büthe. Forthcoming. "Transparency in Qualitative Research: An Overview of Key Findings and Implications." *Perspectives on Politics*.

Kapiszewski, Diana, Lauren MacLean, and Benjamin L. Read. 2015. *Field Research in Political Science: Practices and Principles*. New York: Cambridge University Press.

_____. n.d. "Dynamic Design: Iteration in Field-Based Research." Unpublished manuscript.

MacLean, Lauren M. 2013. "The Power of the Interviewer." In *Interview Research in Political Science,* edited by Layna Mosley, 67-83. Ithaca: Cornell University Press.

MacLean, Lauren, Elliot Posner, Susan Thompson, and Elisabeth Wood. Forthcoming. "Research Ethics and Human Subjects: A Reflexive Openness Approach." *Perspectives on Politics*.

Monroe, Kristen Renwick, ed. 2005. *Perestroika!: The Raucous Rebellion in Political Science*. New Haven: Yale University Press.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

Yanow, Dvora. 2006. "Neither Rigorous Nor Objective?: Interrogating Criteria for Knowledge Claims in Interpretive Science." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, edited by Dvora Yanow and Peregrine Schwartz-Shea, 67 – 88. Armonk: M.E. Sharpe.

Veil of ignorance Process Tracing and Contested Critical Observations

James Mahoney
Northwestern University

The problem of bias in the sources and evidence used to evaluate explanations in qualitative research is a longstanding topic (Goldthorpe 1991; Lustick 1996; Isacoff 2005; see also Thies 2002). The specific problem addressed by Copestake, Goertz, and Haggard's Veil of ignorance Process Tracing (VPT) does not concern biases built into the available evidence itself (e.g., systematic error across all sources). Rather, Copestake, Goertz, and Haggard are concerned with investigator-induced biases in the *selection of sources* and the *selection of specific pieces of evidence* from within a given source. As a solution to these selection problems, the authors propose a new method—Veil of ignorance Process Tracing—in which a "veiled" research assistant is used to help neutrally select sources and evidence as well as code and interpret that evidence.

How serious of a problem is "cherry picking" sources and evidence in case study and small-N research? I think the extent of this problem should not be exaggerated. Scholars often explicitly acknowledge that the existing literature presents competing interpretations of events. In the course of their case narratives, researchers frequently acknowledge differences among historians and historical experts in the interpretation of particular events or processes. Mahoney and Villegas (2006) argue that comparative-historical researchers routinely call attention to major differences in the historiography if these differences are important for their arguments. The professional penalty of ignoring or downplaying obvious contrary evidence is to have one's scholarship regarded as misleading, poorly informed, and not worthy of publication in peer-reviewed outlets.

Researchers often discuss how they work to resolve differences among sources. They may follow the contemporary consensus position in the literature, drawing on more recent studies that highlight flaws in earlier interpretations. They may side with the interpretation that is grounded in the more careful and thorough research. They also often explore the implications of a particular contradiction in the evidence for the specific argument being advanced. Researchers often try to build their arguments by using "basic information" —information about well-known events that is relatively free of interpretation and not subject to a high level of contestation (Thies 2002, 353-354). Quite frequently, they discover that divisions within the literature do not have a bearing on their arguments. To provide an example from my work on colonialism in Spanish America (Mahoney 2010), historical demographers disagree quite a lot about the size of the indigenous population immediately prior to colonialism (e.g., did 5 million or 20 million people live in Central Mexico in 1492?). Given my argument, however, I was

able to note these differences without having to stake out a specific position one way or the other.

My view is that "cherry picking" is a major threat to validity for only select inferences in case study and small-N research. As such, I do not believe that Veil of ignorance Process Tracing is needed for all aspects of process tracing. However, I do see Veil of ignorance Process Tracing as offering an interesting solution to a particular and serious problem that often arises in case study and small-N research. This problem is the use of *contested critical observations* as the basis for reaching conclusions about the validity of a theory or proposition. VPT is a potentially powerful tool for dealing with contested critical observations in case study and small-N research.

Contested critical observations (CCOs) are at the intersection of critical observations and contested observations. A *critical observation* is evidence or data that is highly consequential in the evaluation of a proposition or theory. In the context of process tracing, a critical observation is a piece of evidence or a set of evidence that substantially alters prior beliefs about the validity of a proposition or theory (Barrenechea and Mahoney 2019; Fairfield and Charman 2019). A *contested observation* is a piece of evidence or a set of evidence (including perhaps an entire source) whose validity is disputed. The nature of this dispute may involve either descriptive or causal inference, and the dispute may or may not be a highly politicized and normatively charged.

The methodology developed by Copestake, Goertz, and Haggard offers a useful strategy for tackling the problem of contested critical observations. At an early stage of research, a veiled assistant could help the principal investigator (PI) determine whether specific observations are in fact CCOs that need to be carefully handled. Under this approach, the PI would direct the veiled assistant to various observations that could be important and contested. The advantage of having a veiled assistant at this stage is that she or he can remain neutral with respect to the question of whether these observations are critical and whether they are highly disputed. For qualitative researchers, it can be quite helpful to have a neutral viewpoint regarding the extent to which a critical observation is contested in the literature. The PI may be more apt to view a helpful CCO as less contested than it really is. Likewise, the PI may be more apt to view an unhelpful CCO as more contested than it really is. The use of a veiled assistant could help assure both the PI and the broader scholarly community that key observations have been correctly weighted and appropriately discussed when contested.

Figure 1 presents a two-by-two table in which observations are situated along two dimensions: criticalness and contestation. My argument is that Veil of ignorance Process Tracing is mainly useful for inferences that depend on observations from cell 2, where both criticalness and contestation are high. If critical observations are not highly contested (cell 4), VPT is not needed to ensure unbiased interpretation. If contested observations are not critical (cell 1), the payoff of VPT is probably not worth the effort, time, and resources. I am arguing that VPT is a methodology for dealing with inferences that depend on observations that are both very influential and highly contested—that is, CCOs.

Figure 1
Types of Observations

		CRITICALNESS	
		Low	High
CONTESTATION	**High**	*Contested Routine Observations* **Cell 1**	*Contested Critical Observations (CCOs)* **Cell 2**
	Low	*Routine Observations* **Cell 3**	*Critical Observations* **Cell 4**

Allow me to discuss the example of late-colonial Chile from my own work on colonialism and development in Spanish America (Mahoney 2010, 174-76). One part of the argument about Chile turned on the question of whether this region was a marginal colonial territory during the late Bourbon period of Spanish colonialism (from about 1770 to 1820). I worked to gather all the secondary sources I could find relevant to this issue, and I concluded that the bulk of the evidence suggested that Chile was a colonial periphery. However, a few pieces of evidence and one particular source suggested that it was actually a vibrant colonial region. I noted in the text that some evidence was consistent with an alternative view, and I briefly discussed the contradicting source in a footnote (331). However, for this specific set of observations, I believe that VPT would have been an attractive approach. It would have helped resolve what I still consider to be an important question regarding my argument: Would a neutral (though knowledgeable) reader looking at this set of evidence arrive at the conclusion that Chile was a late-colonial periphery?

Another example is Jack Levy's (1990-91; 2007) work on the origins of the continental war in Europe that led to World War I (see the summary in Copestake, Goertz, and Haggard's Figure 2, this issue). How might VPT

have been useful to Levy in developing and assessing explanations of the continental war? It seems doubtful that a veiled assistant would have been helpful with the basic theoretical framing of Levy's explanation, which views the war as a product of Russian and German mobilization into the local war between Austria and Serbia. Nor would a veiled assistant have been helpful for informing Levy about the kinds of factors that are normally emphasized in explanations of wars in general and the continental war in particular. Levy knows as much as anyone about these topics.

Instead, the assistant could help in two other ways. First, the assistant could help decide whether any observations used by Levy are contested critical observations. For instance, consider the causal chain (circles 1-6 in the summary figure) leading to the local war between Austria and Serbia. A veiled assistant could examine these individual linkages and identify which ones are clearly supported by the literature and which ones seem more contestable. This information can be shared with Levy, who could then report the disputes in an appendix or even better using an active citation system (Moravcsik 2014). Second, for these contested linkages, the assistant can help identify the alternative factors and explanations that Levy may want to consider. For instance, what other factors besides those listed (circles 1-4) tipped the balance in favor of the Austrians initiating war with Serbia (circle 5)? Are any of these additional factors more important than those listed by Levy? At the end of the day, Levy's original explanation may not change. Nevertheless, Levy could report that a veiled assistant reviewed the individual linkages in the chain, looking for any obvious discrepancies in the sources and evidence.

What about selecting the sources to use when conducting research on a given topic? Is VPT useful for avoiding biases with that task? With case study and small-N research, my view is that a veiled assistant would generally offer only modest value-added in the selection of sources. The main reason why is that excellent case study and small-N research requires that the investigator her or himself become an expert in the literature for every one of the cases under analysis. Developing this expertise often requires years of reading and studying the literature and sources related to the cases under study. I think it would signal serious problems if a novice research assistant could come along and point out important additional sources of which the principal investigator was unaware.

The place where the veiled assistant could be helpful is identifying sources for secondary cases in small-N research. These cases receive only brief treatment because

they are used for strategic purposes in comparative research, such as introducing a control case when evaluating a particular causal factor. Again I see the role of the assistant as targeted: The assistant would gather sources relevant to the strategic way in which the researcher is using the case. This kind of research assistance is already often used by scholars. Professors often hire RAs to help with gathering literature and evidence for secondary cases. The difference with VPT is that the assistant must not know in advance what kinds of evidence will and will not support the PI's hypotheses. Hence, the assistant has every incentive to be as neutral as possible when identifying sources.

Perhaps the main context in which VPT can be useful for selecting sources is medium-N studies (e.g., 20 to 100 cases). With these studies, the principal investigator usually cannot be an expert on every single case under analysis. Moreover, the goal of the PI is not to develop a rich sequential narrative for each case. Instead, the PI is interested in specific propositions (e.g., a specific causal linkage) and the observations that do and not support their validity. For instance, Haggard and Kaufman (2012) test theories proposing that inequality shapes democratization via the mechanism of distributional conflict. They code a medium number of cases across a small number of variables, including especially the variable of distributional conflict. Would an unbiased researcher arrive at the same codes for all cases? Certainly, asking and answering this question is good social science. I think the veiled coding of particular variables often makes sense when conducting medium-N research. With high-quality qualitative research, such as Haggard and Kaufman (2012), the investigators have worked hard to put the veil over themselves throughout the research process. Well-trained qualitative researchers repeatedly ask themselves how an objective critic might find and use evidence to show that their argument is wrong. That said, good scientific practice suggests that actually putting the veil on from the start is far better than trying to imagine wearing a veil after you have already seen much evidence.

Practical issues of implementation would doubtless stand as obstacles to using even the focused kind of VPT that I have discussed. Perhaps most important, the new central actor in this methodology—the veiled research assistant—would need to have appropriate characteristics to help carry out high-quality VPT. For instance, let us pretend that we seek to reanalyze Haggard and Kaufman (2012) using a veiled assistant. What kind of assistant would be most appropriate? To begin, we would need an assistant who either does not know this Haggard and

Kaufman article or who does not know we are reconsidering its findings. The assistant probably would need to have a background in comparative politics and excellent overall research skills. The assistant would also need a great deal of time and energy to go through every single one of Haggard and Kaurman's cases! It seems unrealistic to imagine that a single research assistant could do a good job carefully coding every single case using an unbiased sample of sources.

Instead, what I have argued in this paper is that one might use VPT to deal with the contested critical observations (CCOs) in Haggard and Kaufman's (2012) article. In this particular article, the CCOs would be particular codes of distributional conflict that are influential but potentially wrong, given all of the evidence. The challenge would be identifying which specific codes fall into the CCO category. Haggard and Kaufman themselves explicitly note some borderline cases, indicating that they coded them conservatively to work in favor of the theories they critically assess. Nevertheless, one might still want to direct a neutral veiled assistant to explore these cases more carefully.

Methodologists could also develop general rules and heuristics for veiled assistants to use when locating potential CCOs. When deciding if an observation is a CCO, one can start by asking about the extent to which sources agree and how many sources agree or disagree concerning the truth of the observation. What is the nature of the disagreement? One could then follow guidelines for weighting different sources and their interpretations. Some of the factors considered might include the year of publication of the source; the quality of the press, journal, or outlet that published the source; the extent to which the authors of the source have potential biases; other author characteristics, such as reputation, job rank, and employer; the extent to which the source's evidence is directly on topic; and the kind and quality of the primary sources that are used to derive conclusions. In developing rules, researchers could certainly draw on insights from historians, who have thought a lot about how to deal with conflicting sources and interpretations. Through the development of more objective rules for identifying contested critical observations, scholars could make the most of a veiled assistant's time and skills.

I want to end on a note of thanks to the authors for helping us think through the ways in which qualitative research might be enhanced by introducing neutral assistants who do not have a stake in the results. Whether or not Veil of ignorance Process Tracing becomes a significant new instrument in the qualitative toolkit, Copestake, Goertz, and Haggard have done a service by stimulating this discussion about the role of researcher biases in process tracing and qualitative analysis. Their work encourages us to repeatedly ask and seek answers to the following question: What would an objective investigator find?

References

Barrenechea, Rodrigo, and James Mahoney. 2019. "A Set-Theoretic Approach to Bayesian Process Tracing." *Sociological Methods and Research* 48, no. 3 (August): 451-484.

Fairfield, Tasha and Andrew Charman. 2019. "A Dialogue with the Data: The Bayesian Foundations of Iterative Research in Qualitative Social Science." *Perspectives on Politics* 17, no. 1 (March): 154-167.

Goldthorpe, John H. 1991. "The Uses of History in Sociology: Reflections on Some Recent Tendencies." *British Journal of Sociology* 42, no. 2 (June): 211-30.

Haggard, Stephan, and Robert R. Kaufman. 2012. "Inequality and Regime Change: Democratic Transitions and the Stability of Democratic Rule." *American Political Science Review* 106, no. 3 (August): 495-516.

Isacoff, Jonathan B. 2005. "Writing the Arab-Israeli Conflict: Historical Bias and the Use of History in Political Science." *Perspectives on Politics* 3, no. 1 (March): 71-88.

Levy, Jack S. 1990/91. "Preferences, Constraints, and Choices in July 1914." *International Security* 15, no. 3 (Winter): 151-186.

Levy, Jack S. 2007. "The Role of Necessary Conditions in the Outbreak of World War I." In *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, edited by Gary Goertz and Jack S. Levy, 47-84. London: Routledge.

Lustick, Ian. 1996. "History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias." *American Political Science Review* 90, no. 3 (September): 605-618.

Mahoney, James. 2010. *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. New York: Cambridge University Press.

Mahoney, James, and Celso M. Villegas. 2007. "Historical Enquiry and Comparative Politics." In *Oxford Handbook of Comparative Politics,* edited by Carles Boix and Susan C. Stokes, 73-89. Oxford: Oxford University Press.

Morvacsik, Andrew. 2014. "Transparency: The Revolution in Qualitative Research." *PS: Political Science and Politics* 47, no. 1 (January): 48-53.

Thies, Cameron G. 2002. "A Pragmatic Guide to Qualitative Historical Analysis in the Study of International Relations." *International Studies Perspectives* 3, no. 4 (November): 351-372.

The Veil of ignorance Process Tracing (VPT) Methodology: Some Practical Considerations and Limitations (to Decoupling)

Sandrino Smeets
Radboud University

With the introduction of their new approach to process tracing (PT), Copestake, Goertz, and Haggard seek to tackle what is arguably the most prevalent and most serious shortcoming in qualitative case study research: the "cherry-picking" of cases, data sources, or evidence to fit the claim or theory that the researcher is trying to substantiate. The authors should also be complimented for their attempt to make process tracing analyses "more rigorous and transparent" (this issue). By now, we have a variety of handbooks that offer high quality, hands-on advice on how to set up and conduct a process tracing analysis (see e.g., Beach and Pedersen 2019; Bennett and Checkel 2014). However, the term "process tracing" is still being stretched by scholars, who essentially provide us with a narrative reconstruction of a (causal) process. The question "how valid are my causal inferences compared to the inferences of others?" remains central to process tracing analyses. And we should welcome attempts to further develop the methodology of *causal attribution*, specifically for within-case analyses, in which we deal with one or a few instances of a cause-and-effect relationship.

The proof of the pudding, however, is in the eating. So, the relevant question is: Would I be able to apply the VPT methodology in my own research and, more importantly, would it lead to better causal process analyses? One central causal claim in my research concerns the important role and influence of institutional actors—within the context of the European Union (EU) —vis-à-vis national actors in determining the course and outcome of international negotiations (see e.g., Smeets and Beach 2020). Within the field of EU studies, the most influential actors are generally considered to be Germany and, to a lesser extent, France (see e.g., Moravcsik 1998). If I were to go against the dominant interpretations that it is "all about Germany," and seek to show that institution-

al actors, like the European Commission and European Council President, play a prominent role, would VPT help me do a better job?

The solution that Copestake, Goertz, and Haggard offer to the problem of cherry-picking is straightforward: to *decouple* the process of collecting and recording/coding the raw data (exploratory phase) from the process of interpreting/evaluating that data in relation to theorized causal mechanism (confirmatory phase). The exploratory work is to be done by a competent, but not too informed or involved, research assistant (RA), while the confirmatory work is left for the principal researcher (PR). To be able to conduct the fieldwork, the RA should be somewhat familiar with the field but not too informed about the PR's theoretical predispositions. This would allow the RA to collect "pure," "untainted" empirical material regarding the role and influence of particular actors—what Copestake, Goertz, and Haggard refer to as "causal claim Lego blocks" (this issue). The PR would subsequently receive this complete set of blocks of causal process observations (CPO), and therefore not be able—or less inclined—to focus only on the blocks that support her theoretical point of view.

There is some merit to the idea of "blindfolding" the data gatherer. Confirmation bias is a persistent problem, specifically (but not only) in the social sciences, in which definitive proof in favor or against a particular theorized hypothesis is often hard to come by. In the field of EU studies, any scholar will be able to find supporting evidence for the claim that either Germany or the European Commission was very influential in the negotiations, particularly when interviewing stakeholders from one of these two sides. In fact, even a view that Malta has been very influential in EU negotiations can be grounded in *some* empirical observations. The question is whether this evidence, and the interpretation of this evidence, holds