Large-N Qualitative Analysis as an Explanatory Approach to the Medieval Islamic Paper Revolution

Bosco Garcia

Part I: LNQA in theory

1. Introduction

In the first chapter, I considered Pitt's (2001) dilemma of case studies. There, he argued that case studies are impossible in the history of science. Case-studies researchers, he claims, can only proceed from the particular to the general, or from the general to the particular. if we go from the general to the particular, we might be biased in our selection of cases; in other words, we run the risk of cherry-picking. Things do not look much better for the second horn, because—he claims—we cannot make general claims about an entire population from knowledge of a single unit, "or even two or three" (Pitt 2001, 373). In this chapter, I want to challenge the second horn of the dilemma by showing that, contrary to what Pitt and others believe, case studies are able to extract relevant generalizations in the history of science. In a later chapter, I will address the second horn more thoroughly.

Part of Pitt's assumptions about case studies can be addressed, I believe, by considering how methods from the social sciences do extract relevant generalizations from case study work. In fact, I think it is an unfortunate fact of most of the recent work in HPS, particularly the work that focuses on the relevance of case studies in the history of science, that there is no engagement with how the social sciences generate general claims from the historical record. (For an overview of some of these debates, including one chapter that engages with work in the social sciences, see Tilman and Raphael (2016) and Pietsch (2016), respectively.) To remedy this, I will present a recent method from the social sciences that uses case studies to extract causal generalizations—Large N Qualitative Analysis (LNQA)—and show how we can leverage this method to extract causal generalizations from historical cases.

The LNQA methodology, I think can be argued, arises as a solution to two important problems that are traditionally associated with case study research. Such arguments against case studies assert, first and in line with Pitt, that it is not possible to obtain any robust, general knowledge about a population of phenomena when only observing one unit in the population. In other words, case study research does not get off the ground because it cannot produce knowledge beyond its unit of analysis.

This is made even more problematic if we consider that, in many cases, the knowledge we want to obtain is causal in character. It is easy to obtain correlations, both in

history and beyond. The challenge is turning these correlations into claims about causation. We may want to know, for example, whether durable authoritarianism requires a previous social revolution (as Goertz and Haggard often use as an example, as we will see). We may also want to know, closer to our area of interest here, the extent to which changes in the communication infrastructure reliably produce changes in scientific practices (as I will consider later on). The problem is that—so the argument goes—knowledge about a single unit, deep as it may be, cannot tell us anything about the causal structure of the population that the unit belongs to.

LNQA addresses both problems by a "brute force" method. That is, it designs a methodology that allows us to obtain knowledge about general causal claims, such as "If social revolution, then durable authoritarianism", via the study of *all* the cases that fall within the scope of the causal claim. The "Large N" part of LNQA stems from the fact that it considers more than one case to establish the causal claim, and the "Qualitative Analysis" part refers to the reliance of qualitative, "thick" description of the cases within the scope. The main task, then, is to show exactly how to combine both legs: Dealing with many cases while still obtaining knowledge about how general causal mechanisms operate in them. This is done by setting clear scope conditions that apply to a broadly applicable claim (Section 2), and leveraging potential mechanisms that are relevant—and, of course, present—in the cases within the scope (Section 3).

In the first part of this chapter, I want to describe in detail Goertz and Haggard's methodology for the study of causal claims in the social sciences as they are described in *Large-N Qualitative Analysis: Causal Generalizations and Multimethod Research* (2025). I choose their recent account because it presents a systematic view on the methodology, although the methodology itself has been practices in different forms for at least a decade. The point I will defend there is simple: The objective of Large-N Qualitative Analysis (LNQA) is to support causal generalizations in the social sciences. I will make this idea precise, as well as lay out exactly what steps LNQA takes to do this.

An important step in Goertz and Haggard's methodology is delineating the scope of the causal claim. Once a scope is established, LNQA proceeds to make a claim about a regularity that can be subject to causal analysis. A discussion of such regularity claims will be the object of Section 2. Only once a causal mechanism has been demonstrated to operate in the cases, however, can a regularity be a generalization proper. LNQA chooses to use causal diagrams in order to capture the mechanism responsible for turning regularities into generalizations. In Section 3, I will go back to case studies and describe the interplay between cases and causal generalizations, specifically how the cases within the scope of the generalization should show, up to a point, the causal mechanism operating in them. To

do this, I will also discuss Goertz and Haggard's use of causal diagrams and their role in LNQA. In Section 4 I take up the question of "up to what point" the cases within the scope of the claim should express the proposed mechanism in operation.

In the second part of this chapter, I will run a translation exercise, and show how an LNQA approach to my field of interest in the history of science would work. The point there will be illustrative, and thus some important details that would be imperative in an actual LNQA study will be missing. However, I think that the methodology is particularly apt for causal claims in the context of the history of scientific practice, so I will sketch how it could be helpful for case study work in the context of material culture in the medieval Islamic world. Specifically, I will consider the problem of how the introduction of paper in the Islamicate world contributed to the scientific culture of the period. As I will show, this is a problem ripe for LNQA study, as the historiography has up to now only provided reasons to believe an association between both processes, without an understanding of how one can cause the other.

I will reproduce the structure of Part I in the second Part, with the purpose of showing exactly how we can put GH's theory to use. In Section 5, I will spend some time getting clear on exactly what regularity we should use to start our investigation. As I will explain, this is a crucial part of the research process, and we will need to do some work before we can formulate a clear regularity that is amenable for causal modeling. Finally, in Section 6 I will consider the causal hypothesis proper, offering a causal model that explains the regularity, and running some within-case inference to argue for the validity of the model. In the conclusion to Part II, I will extract some conclusions and tie back the discussion to the question about the role of case studies in the history of science.

2. Establishing regularities

The aim of this section is to get a clear explanation of what a "regularity" is in GH's framework. Regularities are a kind of association. Specifically, they are "associations which may be causal but for which the causal relationship is not established by the combination of theory and within-case causal inference" (43)[1]. The fact that they are associations is important, because that is what sets them apart from causal generalizations, the main object of consideration in LNQA. Causal generalizations are regularities that are convincingly explained in the cases within its scope by a causal mechanism.

Since we are interested in obtaining causal knowledge, the form of regularities (and generalizations, too) is usually of the form: *If X* (the cause)*, then Y* (the effect). Note that, for

[1] Whenever I am citing Goertz and Haggard (2025), I will simply include the page number in parentheses. Whenever I am citing other sources, I will include the year of publication and page number.

now, we are not interested in what kind of relationship mediates X and Y. So far we may understand the relationship between X and Y purely associationally: Whenever we observe X, Y tends to be present as well. We want such a statement to be regular, in the sense that observing X reliably leads to the observation of Y.

Establishing a clear scope of a regularity is crucial here, because without it we cannot really know whether the claim is truly regular or not. Imagine that we have X = Unemployment, and Y = Poverty. We want to establish that being in poverty is regularly associated with being unemployed. In order to do this, however, we need to know exactly what we mean by "unemployed" and by "poverty". Do underemployed individuals (for example, working less than twenty hours a week) count as unemployed? Do we include a smartphone in the list of things that a standard, modern individual needs to live in modern society?

In the context of LNQA, we are interested in the "cases" that may fall within the scope of X and Y. So we may think about the aforementioned claim in national terms, and think about each nation as cases: If the rate of unemployment (given a convention of what we mean by the term) reaches a certain level, is that associated with a certain poverty rate (again given a convention for the term)?

In the literature, the word "scope" is sometimes related to the bounds within which a claim is taken to be valid. In the above example, the scope of the claim would be the cases were unemployment would be actually associated with poverty. GH's understanding of the term is different, and not directly related with the cases that do conform to the regularity. The scope is the "set of cases determined by the proposed generalization or regularity" (18). In this context, it would be the entirety of the world's nations, because all nations experience a level of unemployment and a level of poverty. Some might conform to the regularity—meaning that the correlation will obtain—or not—hence that it does not obtain––but testing conformity is something we do *after* establishing the scope.

LNQA is characteristic in requiring that *all* cases within the scope of the claim are considered to establish how regular the claim is. This is where the methodology leverages "brute-forcing" to establish the regularity (and eventually the generalization) of its claims. "Since all cases in the scope are subject to scrutiny", GH put it, "concerns about the representativeness are eliminated altogether" (18). Indeed, if we consider literally all cases within the scope of the claim, we do not even need to ponder anything about sampling, since we are extracting regularities about the entire population.

The trade-off here is that we exchange considerations about probabilities to considerations about the "strength" of the claims within its associated scope. I will say

more on this in Section 4, but in LNQA the strength of a regularity or a causal generalization is measured by the percentage of cases out of the total scope that conform to them. This applies both to regularities and causal generalizations. Is there a difference, then, between a strong regularity and a strong causal generalization? The difference lies in the standard to show that a case conforms to the purported regularity or generalization. As we will extensively see later on, regularities are associations, whereas causal generalizations are regularities backed up by a causal model and within-case inference. To show that a case conforms to a causal generalization, one needs to show that both X and Y obtain, *and* that they obtain in the case in the way implied by the causal model. A regularity and a generalization of the same strength, then, though equally "strong", are not the same, because the standard of conformity for the generalization is much more demanding.

Coming back to scope, the reader might feel dizzy imagining running a case study for each and every one of the nation-states of the world in order to test the claim above. That reaction is understandable. Since LNQA requires that we deal with all cases within a scope, we might want to restrict our analysis to "rare" cases—specifically, cases where the Y condition is rare. But here there are two important caveats to consider. First of all, as GH point out, there is no such thing as an "intrinsically" rare event, and what we consider "rare" will depend on things such as the intension of the concept, which affects the number of cases it refers to. (22) I will touch on this later on in this section, but one might for example think of poverty as extreme poverty, thus reducing the number of countries that satisfy the Y condition.

Finally, it is important to note that the constraints in dealing with all cases within the scope of a claim "are not necessarily in the sheer number of cases, but in the complexity of the causal argument and the nature of the evidence that must be brought to bear" (18). This is a part of the LNQA method that will be most present when we try to turn a regularity into a causal generalization. In addition, it may not be necessary to deal carefully with every single one of the cases to produce strong generalizations. GH mention the example of Rosemary Kelanic's (2012) work on the oil strategies of great powers, who omitted Russia and France out of 14 possible cases (leaving four total nations) and nevertheless found that 10 out of 11 conformed with her causal hypothesis, thus having a strong causal generalization.

So, now we know that the scope of a regularity is all the cases determined by the regularity. But, if this is the case, then the form in which the regularity is obtained will be relevant for the researcher, because it will influence how she deals with the cases and which cases are appropriate to consider to establish the causal strength of the claim. GH

consider three types of regularities: X regularities, Y regularities and X-Y regularities. I will briefly explain each of them in turn.

X regularities (or sufficient condition claims) establish that *if X, then Y*—i.e., that some phenomenon of interest with characteristic Y follows the presence of phenomenon with characteristic X. GH operate with these variables in a Boolean way, such that X,Y=0 represents the absence of the condition and X,Y=1 its presence. An X regularity, then, is "the percentage of cases of X=1 that are also Y=1" (46). An example of X regularities is Way and Levitsky's (2022) "durable authoritarianism" claim, according to which social revolutions consistently lead to durable authoritarian systems. In GH's framework, "if social revolution [X], then durable authoritarianism [Y]" (26).

An important feature of the way that GH think about X and Y, or the purported features of the causal claim with their associated scope, is that they are amenable to analysis in two-way tables. (Figure 1) The columns codify whether the cause is present or absent, and the rows whether the effect is present or absent. In this way, we can extract the percentage of cases of interest for any combination of values of X and Y. In Figure 2, I present the two-way table created by GH to present the results of Levitsky and Way (2022). From the table, we can read that the percentage of cases where both X and Y are 1 is 0.78, which as we'll discuss later is considered by GH to be a strong regularity, since more than 75% of cases conform to it.

	$X = 0$	$X = 1$
$Y = 1$	—	Causal mechanism cases
$Y = 0$	—	Nonconforming cases

	$X = 0$	$X = 1$
$Y = 1$	Nonconforming cases	Causal mechanism cases
$Y = 0$	—	—

Figure 1. *Left: Two-way table for X regularities. X regularities. In X regularities, we are concerned with the right column of the two-way table, where the cause is present (expressed as X=1). Right: Two-way table for Y regularities. In Y regularities, on the other hand, we are interested in cases where the effect is present, meaning Y=1. This happens in the top row of the table.*

	No-revolution	Revolution
Durable authoritarianism	227	14
Not–durable authoritarianism	110	4

Figure 2. *Results from Way and Levitsky's (2022) analysis of durable authoritarianism. X = Presence or absence of social revolution, and Y = Presence or absence of durable authoritarianism. Note that two revolutions are not included in the dataset.*

Y regularities (or necessary condition claims) establish that *if Y, then X*. Since *Y* denotes the effect that applies to the cases within the scope, the claim is the inverse. While in X-generalizations we start with some condition and consider the strength to which an effect is associated with it, in Y-generalizations we run the inverse procedure: We start with some effect and consider what all cases that instantiate it have in common. This sets the

necessary conditions of the effect, to the extent that such condition is necessarily present for the effect to obtain, even if they may not guarantee its occurrence. Thus, Y-generalizations are "the percentage of cases where Y=1 where … X=1" (p. 43).

GH consider Deyo's (2020) hypothesis explaining the economic growth of East Asian societies during the later part of the 20th century; however they do this using Geddes' (2010) systematic data, which they condense into a table that I reproduce here. (Figure 3) Deyo proposed in 1989 that one of the main reasons why countries in East Asia experienced extraordinary rates of economic growth (what was to be called the "East Asian miracle", (Birdsall et al. 1993)) was high levels of "labor repression", meaning very weak labor movements. While Deyo selected four high-growth East Asian countries—South Korea, Taiwan, Hong Kong and Singapore—Geddes took on testing this hypothesis in a more general, upper-income population of countries.

	Low labor repression	High labor repression
High economic growth	0	6
Not–high economic growth	16	12

Figure 3. *Geddes' (2010) systematic review of 34 East Asian Countries, condensed by GH (p. 44). X = Presence or absence of high economic growth (a cutoff of 5% growth is imposed), and Y = High or low labor repression.*

For Y regularities, we need to look at the top row of the table: Given some effect (High economic growth), what proportion of cases feature high labor repression? Looking at the table, we can see that the totality of those cases fulfills that condition. Of course, we can also see a much higher number of lower growth countries that also experienced high labor repression, but this just shows that high labor repression is a *necessary*, not sufficient, condition for high growth. GH gather from Geddes' analysis that Deyo's hypothesis was ultimately vindicated: "the East Asian newly industrializing countries constituted a distinctive export-oriented growth type in which labor subordination mattered" (p. 45).

For completeness, I will also briefly explain X-Y regularities. If X regularities represent sufficient conditions, and Y regularities necessary conditions, then X-Y regularities are *necessary and sufficient conditions*. In tables such as those in Figures 1-3, strong X-Y regularities are those in which at least about 75% of the observations fall on the diagonal (the (1,1) and (0,0) cells), i.e. where both X and Y are present or both are absent. GH make the point, however, that it is "relatively unusual" to find cases amenable to XY-generalization (but see p. 48-50 for an example).

Now, while Deyo's example contains no non-conforming cases, we can see that there are 4 non-conforming cases in Way and Levitsky's table (Figure 2), where even though the cause was present (there was a social revolution) there was no durable authoritarianism

that followed. GH spend some space discussing what strategies are available to LNQA researchers for dealing with such cases. Before concluding this section, I will briefly explain what those strategies are, since they can further illuminate what regularities are and how they relate to their scope, as well as their purported causal mechanisms, that I will discuss in a moment.

The first strategy is to narrow the scope of the regularity. The idea here, to put it simply, is to decrease the extension of a concept by increasing its intension (this formulation comes from Goertz (2020)). In other words, we reduce the scope of a concept by increasing its attributes, so that fewer cases are able to meet its specifications. GH mention the work of Kurtz (2000) as an example of narrowing scope, who explores distinct definition of the term "peasant" in the context of counting the number of peasant "revolutions" in the 1940s and 50s. Wider definitions, such as a "minimalist" definition of peasants as "rural cultivators" encompass a much higher number of cases. If we increase the intension of the concept and include conditions such as the ownership of the land, then the number of cases would be reduced (see Kurtz, 97-98 and 98n16).

Another strategy is, as GH call it, to increase the complexity of the cause (X) or the effect (Y). Which one we decide to focus on will depend on the kind of regularity we are interested in studying. Suppose we start with an X generalization: "If X, then Y". In this case, we can increase the number of causes that can lead to an outcome, thus also increasing the number of cases that conform to the regularity. Let's go back for a moment to Levitsky and Way (Figure 2). Suppose we change the regularity from "If social revolution, then durable authoritarianism" to "If social revolution (X_1) AND absence of large-scale military intervention (X_2), then durable authoritarianism". If we include condition X_2, the number of cases with a social revolution (Y) goes down from 18 to 12, but all of them satisfy X_1 AND X_2. This X regularity now has a strength of 100%, although the applicable cases (the scope of the claim) has significantly gone down. This is illustrative of the trade-offs that come with strategies of complexity increasing. (See more details in pages 54-5.)

Y generalizations can be made more complex by assuming what GH call "equifinality": the possibility of "multiple paths to an outcome" (56). LNQA translates this idea in its typical Boolean fashion by using the clause OR. As a brief example, consider Miller's (2021) analysis of processes of democratization. Miller is interested in showing that democracies tend to come after some form of "shock", including civil wars, assassinations or a military defeat in an international conflict. GH calculate (see p. 57 for details) a strength of .72 for the claims "If democracy (Y), then shock (X)"; however, if we include alternative conditions in X, we get a higher strength. Miller adds to the shock condition that of "electoral continuity", in recognition that "autocrats frequently lose power due to elections" (57). If we

call the shock condition X_1, and the electoral continuity condition X_2, then we can implement equifinality by stating the regularity "If democracy (Y), then shock (X_1) OR electoral continuity (X_2)". If we consider two paths together, this gives a strength of 91% for the cases. (57)

A note about equifinality. GH consider two separate notions of equifinality. In the first, we have "two or more causal factors that are functional substitutes of each other" (183). Miller's theory fits within this conception. His theory encompasses two possible paths to the same outcome: If democracy (Y), then shock (X1) or electoral continuity (X2). In the context of Miller's theory, this is a necessary condition claim with an additional assumption that each cause is incompossible with the other: If there is democracy, then it must have happened either through shock or an electoral process, but not both.

This is different to another sense of equifinality, where "the two paths reflect competing theories that are effectively being pitted against one another" (183). In this case we have two *competing* theories to account for the same effect, Y. GH illuminate this case with work by Staniland (2021) to illuminate the second meaning. Staniland consider a number of competing theories for how leftist insurgencies might arise in democracies. In one theory, if there is an "incorporation window" of the leftist movements into the democratic system (X1) and the opposition is fragmented (X2), then there is a leftist insurgency. On the other theory, if there is no redistribution towards the bases of the movement (X1) and the democratic regime blocks the left (X2), then there will be a leftist insurgency. He then considers how many cases conform to *both* theories (1 case), to *none* of the two (2), to the second one *only* (2), and finally to the *first* one only (7). They thus conclude that the first theory has more evidence in favor. The crucial difference with respect to equifinality is that here we do not have a theory that includes both scenarios—rather, we have two different theories with opposed causal conditions to explain the effect. In what follows, I will focus on the first kind of equifinality.

Let's summarize before I move on to discuss causal mechanisms in LNQA. In this section, I have provided an account of one of the main building blocks of the methodology: the construction of associational regularities with clearly defined scope conditions. This discussion was meant to show the crucial importance of clear scope conditions in LNQA. Because the scope of the regularity just *are* the cases where it is supposed to be valid, and because the method requires dealing with all cases within the scope, we need to be clear about the scope before we can investigate the cases. Otherwise, we would not be able to tell which cases would be relevant, and thus we would not be able to tell the strength of the regularity.

Finally, these regularities, as we have seen, are merely (or at least so far) purely observational, in the sense that they don't necessarily imply any causal relationship between X and Y. X and Y, then, are dimensions of the cases within the scope of the regularity, that may or may not be dimensions of a causal process responsible for their co-occurrence. This already suggests that, in order to go from regularity to full causal generalization, we will need to show the presence of a causal mechanism in the cases within the scope of the regularity, so that we can explain with the existence of the regularity. In the next section, I will discuss how this is done in LNQA.

3. From mechanisms back to cases

3.1 Causal diagrams

Causal mechanisms, GH put it, "are central to LNQA" because they "structure the case study research and the selection and presentation of evidence that permits within-case causal inference" (67). Indeed, as I have reiterated, the presence of a causal mechanism responsible for the existence of the regularity under investigation is what distinguishes the mere regularity from the causal generalization, the construction and validation of which is the objective of LNQA.

Causal models, particularly causal diagrams, play the role of "structuring" the research and the presentation of the evidence. In fact, GH attribute an absolutely essential role to causal models, to the point of claiming that "if the argument cannot be mapped onto a simple diagrammatic figure, then it is probably not clearly thought out to begin with" (68). So, in this section, I will discuss the role of causal models in LNQA and how they condition the presentation of evidence in the cases within the scope of the to-be generalization. The causal generalization will be validated only once the appropriate mechanism has been credibly shown to be present in the cases within the scope of the regularity claim. So, first, I will explain how GH think LNQA should make causal models, and then discuss how they constrain the presentation of the cases.

GH think causal models should ultimately reflect the way in which X causes Y, given the particular regularity that we have set out to explain. I will use an example to illustrate these prescriptions concretely, but first I want to lay them out. Causality is represented with an arrow in the model, so that X cause Y is represented as X → Y. The X → Y relationship will typically include some mechanism M, which represents "what causes Y to transpire or … what prevents it from occurring" (78). In the Boolean framework of LNQA, we represent the presence or absence of either element of the model as a {0,1} value. So, for example, if we wanted to represent how X causes Y through M, we would assume X=1, M=1, Y=1; but if we wanted to represent how M *prevents* Y from happening we might want to write X=1, M=1,

Y=0. Causal models can thus represent both the mechanism that leads Y to happen or, conversely, that prevent it from happening.

Consider, as an elementary example, the following causal model, created by GH and based on Coleman's (1986) analysis of Weber's *Protestant Ethic*. (Figure 4) Coleman argued that Weber failed to show how "individual orientations combined to produce the structure of economic organization that we call capitalism (if in fact they did in combination produce this effect)" (1986, 1323). His proposal, explicitly expressed in Figure 4, is that the Protestant religious doctrine changed the values of the actors involved, which in turn affected their economic behavior and thus the development of Early Modern European capitalism. As the Figure shows, the religious doctrine is the cause X, which causes the historical outcome of modern capitalism Y, via the "boat" mechanism M that encompasses both a change in values and in economic behavior.

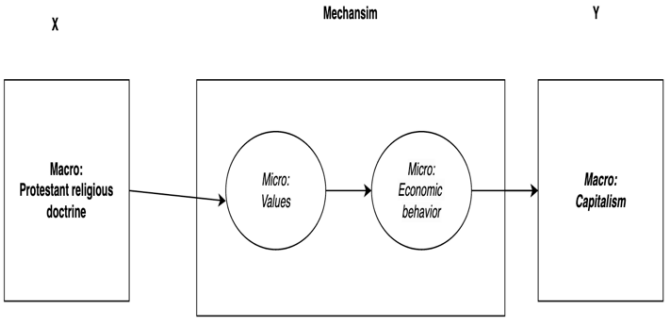

Figure 4. *Coleman's "boat" figure, adapted by GH (73). Weber's argument is reconstructed in an explicit causal manner, leading from the Protestant religious doctrine (X), to the "Spirit" of capitalism (Y). Crucial here is the mediating role of the Values associated with the Protestant doctrine, which would impact economic behavior and thus have an effect on the development of capitalism in the sixteenth century and beyond. The conjoined contribution of both the values and their impact of economic behavior is considered by GH to be the mechanism (the "boat" associated with Coleman that carries both components) mediating cause and effect.*

This example of Weber's *Protestant Ethic* is, of course, elementary, and causal models can be considerably more complex. A causal argument can posit, for example, that one cause can produce the same effect via different mechanisms. If, for instance, we set some CCTV cameras installed in a parking lot to avoid car theft, these can act both by dissuading potential thieves and by redirecting theft to other locations without cameras. In a causal diagram that represents this situation, we would have to include two simultaneous mechanisms. I'll tackle in a little bit how we can do this.

Another crucial requirement to write good causal diagrams for GH is that they should include alternative explanations of the effect. In Figure 5, I include their figure which indicates how these should be included in the causal diagram. These alternative explanations provide a competing cause Q that could be producing the effect, and thus a different mechanism M(Q) that would mediate between Q and Y. Returning to Weber for illustration purposes, one could imagine that the capitalism born in Europe during the sixteenth century might be due to some cause that is not the adoption of Protestantism. We may think with Marx that capitalism (Y) arises from a primitive accumulation (Q), operating through mechanisms M such as colonialism (M_1) or expropriation of the means of production from small land owners (M_2). A robust LNQA explanation of the origin of Early Modern European capitalism à la Figure 4 would also have to consider other potential explanations like Marx's, include them in the causal diagram and discard them after considering the evidence in the appropriate cases.

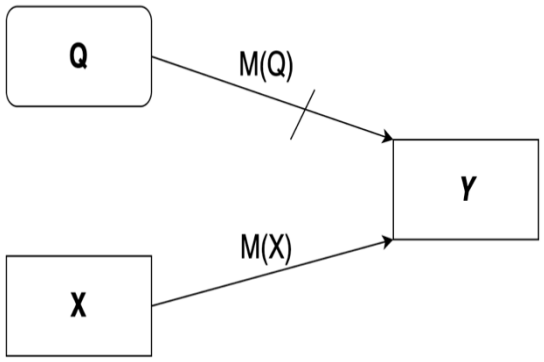

Figure 5. *GH's illustration of how causal models should include alternative explanations (77). Alternative posited causes are represented here with* Q*, and the mechanism through which a cause operates is represented as* M(·)*. We indicate that Q and M(Q) are alternative explanations by crossing the arrow leading to Q.*

Finally, causal diagrams should also include multiple causes that contribute to the same effect, when appropriate. I won't explain these in a lot of detail, but there are two main ways in which causes can be "aggregated", in GH's wording, into the same diagram: vertically and horizontally. In vertical aggregation, "multiple causal factors are stipulated to operate broadly contemporaneously with the outcome variable" (79): that is, we have multiple causes that contribute to the effect, and we take them to more or less do that with some temporal overlap. (Figure 6) We aggregate them "vertically" because LNQA takes causal diagram to represent temporal processes, such that as we move to the right we are "moving the process forward" (my expression). Vertically overlapping causes or mechanisms, then, imply that they are co-occurring in time.

The other type of aggregation is "horizontal aggregation", where "causal processes are modeled as unfolding in time" (79). In this sense, and not surprisingly, horizontal

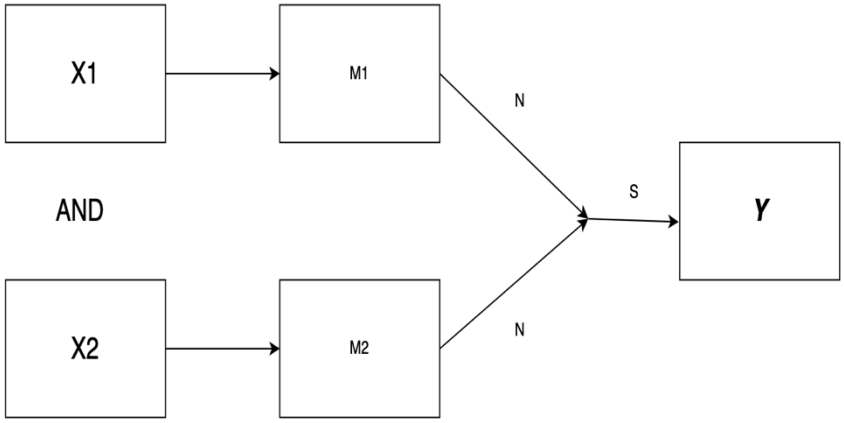

Figure 6. *Vertical aggregation (80). Here, two different causes X1 and X2 are posited to contribute to the effect Y, each with its own associated mechanism. Note that both X1 and X2 are necessary causes to Y (indicated by the N arrows), but they are jointly sufficient (S arrow) to produce Y.*

aggregation is the inverse of vertical aggregation, in the sense that we model precisely the temporal dependence between events that contribute to the causal process. GH offer a typology of at least three different horizontal aggregation strategies. Here, I will just state them and define them before discussing how causal models contribute to the production of causal generalizations in LNQA.

The simplest version of horizontal aggregation are "causal chains", where each step of the causal process takes place in a series, such that in the simplest cases the cause produces some effect through some mechanism, which leads to another further effect, all the way to our effect of interest Y. In another kind of horizontal aggregation, "path dependence" models feature a causal factor "located at some relatively distal time in the past … [with] its own cases and mechanisms" (85). Path dependence models also include a second mechanism which "explains the transmission or persistence of the cause over an extended time period" (85). GH use Homola et al.'s (2020) account of how racist attitudes could persist in Germany long after the Holocaust, who suggest that cognitive dissonance is a mediating mechanism that explains this persistence. Finally, "constraining" causal models feature a cause X1 producing some effect X2, such that X2 actually "blocks" the production of the effect Y. (86-7) The resulting process is one where the presence of X1 leads to the blocking of effect Y. GH mention "deterrence" as an example, with mediating causes such as the possession of atomic bombs preventing the aggression of a rival country.

One final point. Overall, GH think that the creation of causal diagrams in LNQA should follow a "completeness" criterion. What this means is that the method requires "a complete causal graph, a complete set of descriptive inferences from particular historical settings to the graph, and a complete set of inferences about the causal mechanisms that

generate realizations of the causal graph" (97). A complete causal graph should include not only all the relevant causes and mechanisms that lead to the outcome Y, but also the alternative hypotheses that we are discarding on the basis of our evidence.[2]

But this leads to the question of evidence: How do we go from the "particular historical settings" to the graph? What "set of inferences" should we make such that we can make the cases speak to the graph we have created? Without this step, the "loop" (my expression) that characterizes LNQA (from the cases in the scope of a regularity to the cases that provide evidence for the causal story that explains it) is not yet closed. In the next section, I will close this loop by considering how LNQA leverages the cases once a proposed regularity and a causal model have been set up.

3.2 Within-case inference

In this subsection, I will discuss how LNQA makes use of cases to validate (if indeed they do validate) the causal model that purports to explain the generalization. This requires doing within-case causal inference, since the work required at this point consists in showing that the theoretical mechanism we proposed actually operates in the cases. First, I will explain how LNQA carries out causal inference in the abstract, and then I will consider an example in detail that shows how it is done in practice.

Presenting evidence from the cases involve, say GH, three steps: 1. "Matching the concepts with *empirical description*", 2. Providing "evidence that the causal relationships indicated by the arrows are also actually causal in the case", and 3. Aggregating "discrete bits of evidence into an overall assessment of the validity of the claim" (106). In other words, first we have to show that the concepts we have used to formulate the causal model map onto the case, then we consider whether the case displays the mechanism proposed, and finally we consider all the evidence as a whole to see whether the general story matches everything we know about the case. Figure 7 represents this method in a diagram.

A word about aggregating evidence before I show how this is done in practice with an example from the literature. First, it should be noted that the evidence we can use to, say, confirm the presence of X in the case can be extremely varied. GH use Cartwright and Hardie's work to make this point. For them, "empirical evidence" encompasses very different kinds of facts. For example, "Whether the facts are local, like "there is an 8-foot wall around our parking lot," or very general, like the law of gravitational attraction, for us

[2] Of course, it is not reasonable to expect that all causes and mechanisms contributing to an effect can be included in a graph. GH accept that in many cases including them will be unfeasible, or without "theoretical interest" (98). The completeness criterion, then, is relative to both means and relevance in the production of the causal diagram.

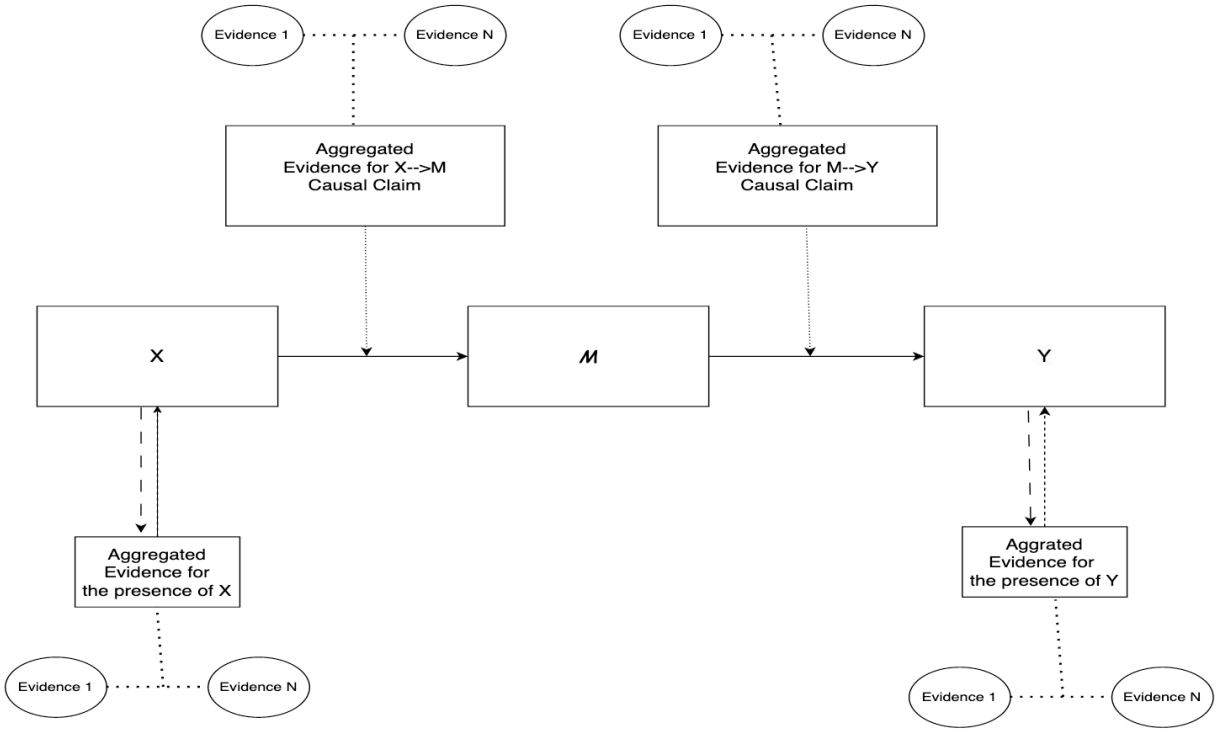

Figure 7. *GH's framework for within-case causal inference (107). In order to tell whether a case supports a causal generalization, first we look for evidence of the presence of the cause X, and evidence for presence of effect Y. To do this, we have to see whether the concepts we have used in our regularity proposition ("If X then Y") have empirical content in the case. If that is the case, the next step is to test whether the proposed mechanism that mediates X and Y operates in the case as is specified in the causal model. If the cause and effect are present in the case, and the mechanism we proposed also operates in the case to produce (or contribute to) Y from X, then we can say that the case supports the causal generalization. In both steps, different kinds of evidence need to be aggregated into an overall assessment of their degree of support. There are many available methods to do this, but GH suggest that process tracing is a natural one for LNQA.*

evidence claims report facts about the world" (2012, 21, cited in Goertz and Haggard (106)). This underscores "the wide variety of evidence that might be brought to bear on a causal proposition" (106).

The question, then, is how to aggregate such variety of evidence into some quantifiable standard that provides a measure of the extent to which we have support for X, X → M, Y or Y → M. GH point out that there are a range of options existing in the literature, including LNQA literature. We can find methods from statistics, specifically OLS, which uses addition and multiplication for the purposes of aggregation (110). But many of the existing LNQA studies draw, "even if not fully recognizing what they are doing", on ideas that stem from the logic of necessary and sufficient conditions. Specifically, they use the van Evera tests, (Van Evera 1997) which are most exploited in the process-tracing literature. (Mahoney 2012) Before illustrating with an example how to put this in practice, let me explain what the van Evera tests are. While they won't show up in the example, I will use them in Part II of this text to argue for the strength of my causal hypothesis there.

The van Evera tests allow the researcher to see how a particular causal hypothesis fares against some evidence that bears on that hypothesis. Imagine, then, that we have one particular causal hypothesis that we want to test. There are four tests, each for one of the 2x2 possibilities that arise from combining necessary and sufficient conditions. That is, one could imagine four possible tests: one that passing it is *sufficient* for the hypothesis to be confirmed (technically, its likelihood greatly increases, as we will see), another that passing it is *necessary* for the hypothesis to be considered viable, a test that is *both* necessary and sufficient for the hypothesis to confirmed, and finally a test that is *neither* necessary *nor* sufficient for confirmation of the hypothesis. I will briefly explain each in turn.

Smoking-gun tests are those which, upon passing them, the hypothesis under consideration substantially increases its likelihood, but failing it only decreases the likelihood slightly—the condition is not necessary, at the end of the day, and so failing the test does not affect the hypothesis catastrophically. What about rival hypotheses that we might be considering? Passing a smoking-gun test substantively weakens those other hypotheses, and failing them only strengthens them a little bit. Key in the context of seeing how passing and failing the test alters the probabilities is that all hypotheses under consideration exhaust all the space of possibilities. As a result, a change in the probability of one implies a correlative change in the rest, if we maintain normalization.

Hoop tests are those which, upon passing them, the relevance of the hypothesis is confirmed, but it does not strengthen it in any substantive way. They are a "hoop" that the hypothesis has to jump through if it wants to be considered as a candidate. As a result, failing the test automatically eliminates the hypothesis. The rival hypotheses, on the other hand, are only slightly weakened or strengthened upon the result of the test.

Following this logic, doubly decisive tests are those that satisfy both necessary and sufficient conditions. They are thus the strongest test that one could run on the hypothesis, in that passing them automatically confirms it and failing them automatically eliminates it. Conversely, if the test is passed, the rival hypotheses are automatically eliminated, and if it is failed the rival hypotheses are substantially strengthened. In Straw-in-the-Wind tests we have the opposite scenario. Since they test for conditions that are neither sufficient nor necessary, passing them only affirms the relevant of the hypothesis, and failing them only slightly lowers the probability of the hypothesis. The rival tests go through the opposite process: if the hypothesis under consideration passes the test, they are only slightly weakened, and if it fails it, they are only slightly weakened.

As I said earlier, now I will discuss a historical study by Deborah Boucoyannis (2021), as GH recount it within their LNQA framework (119-124). Boucoyannis's work sets out to investigate whether the emergence of representative governmental chambers requires the

presence of a previous, strong central state power. This is against previous consensus which assumes, instead, that these representative institutions actually require a weak central authority. Thus, her proposal is expressed in terms of necessary conditions. In her own words, "This study aims to show that strong central powers, especially over the most powerful social groups, are a precondition of representative institutions at least in the major cases typically examined in social science" (2021, 24–25). We can express her regularity, then, as the idea that "high values on X – stronger kings – were more likely to be associated with the emergence of parliament as the central court and organ of governance – the outcome Y" (121).

Her scope, in this case, is the set of states existing in Europe during the Medieval period where we observe the emergence of representative chambers since the creation of the Leonese *Cortes* in 1188. She does not consider each and every case within the scope of the claim, but she does look at a considerable sample of them, specifically the emergence of parliamentary courts in England, France, Castile, Catalonia, Italy and the Low Countries, with some minor comparisons with Russia and the Ottoman Empire in the 16th-17th centuries. As GH do, I will focus on the English case.

Now we have the scope and the regularity. How do we proceed with the causal analysis? Boucoyannis' hypothesis is then one of necessary conditions: *Given* the set of cases that feature parliamentary courts (Y), then those cases must show a strong central power. In other words, she presents a Y regularity that she proceeds to defend is a causal generalization. How does she do this?

For completeness I include GH's causal diagram that stems from Boucoyannis' work. (Figure 8) However, I cannot explain most of the nodes and I will briefly present the key idea about what causal mechanism Boucoyannis suggests that exists between X (strong central power) and Y (representative chambers). The main idea is that representational institutions emerged as the royal power centralized military and judicial functions, and thus increasing sectors of the population claimed a say in the sentencing of justice that was being administered from a single source. A key move in this context is that royal "strength" (the main factor that ultimately contributes to representative government) is operationalized in terms of the accumulation of judicial functions. As these functions are progressively centralized, bigger parts of the population start requesting participation in their administration, eventually forming institutions designed for such participation: the parliamentary courts.

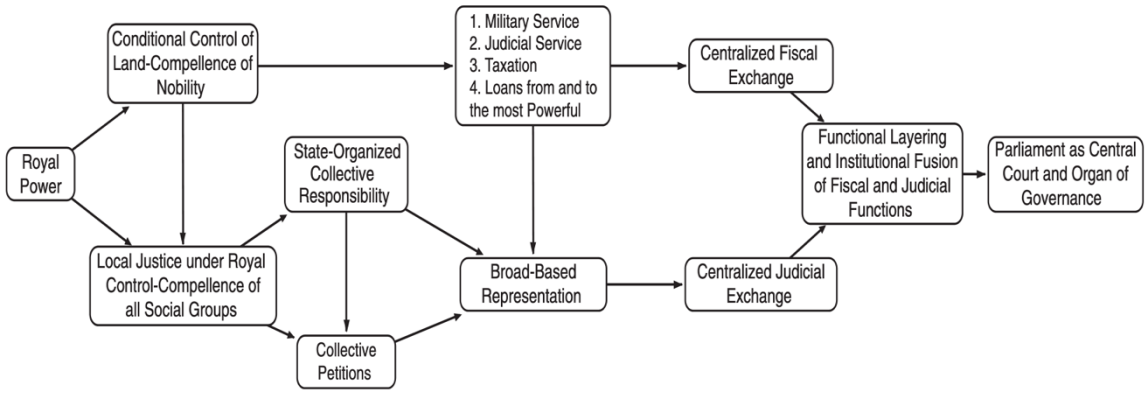

Figure 8. *Boucoyannis's (2021) causal model for the emergence of parliament from royal power (GH, 120). Although the number of discrete mechanisms that contribute to the effect is quite high, the main idea is that, by centralizing the administration of judicial functions, the crown generates an institutional demand for broader participation from broader sectors of the population.*

Boucoyannis then shows that the English crown actually experienced such a concentration in judicial functions, using a variety of indicators (this speaks to the variety of evidence that we saw with Cartwright and Hardie). Just as an example to give an idea of what kind of indicators she considers: The number of hearings that the Crown called substantially increased since the proclamation of the Statute of Westminster in 1275. As a result of such a statute, many "civil, criminal, and administrative matters, from debt to slander, extortion, murder, rape, and the timing of assizes" (Boucoyannis 2021, 49). The royal government became more centralized, then, by promoting regulations that made the Crown the administer of justice in an expanding list of legal violations.

What about the mechanism that mediates between strong royal power and parliamentary institutions? A result of this centralization, more and more sectors of the population had to be physically called to Westminster to attend in hearings. She uses a number of archival documents that testify to this fact. These documents testify to how, initially, these were hearings on a "large scale", where many different cases would be heard. However, by 1290, the role of Parliament was explicitly acknowledged in these documents.

However, Boucoyannis's argument for the role of this mechanism of centralized judicial administration also expands beyond the English case, and she offers a comparative analysis of the English and French crowns. The French state is traditionally considered one of the starkest cases of a strong, centralized government during its ancient regime; however the Estates General (*États généraux*) was one of the weakest parliamentary institutions in the history of Europe, famously meeting extremely intermittently and contributing to the explosive popular revolts of 1789. If this is the case, does the French case disprove the generalization that Boucoyannis is trying to draw? Here, she proceeds to

deny that France actually had strong, centralized royal power during the High Medieval period. In other words, there is no presence of the case, and thus it is no problem to the regularity that we do not observe the effect.

Thus, Boucoyannis uses evidence in her cases to argue for the causal generalization she proposes: Again, that strong royal power is a necessary condition for the emergence of parliamentary institutions in Medieval Europe. In GH's account of her work, she basically fits the description of an LNQA methodology: She has a clear scope of cases, she formulates a clear regularity that she tests in the scope of the claim, and formulates a causal mechanism that is responsible for the regularity that we observe in the cases; additionally, she considers a considerable sample of the cases within the scope of the claim and tests that the mechanism operates as is specified in her theoretical model. She does not make use of the van Evera tests to aggregate the evidence, however, which prevents her from fitting the LNQA framework exactly, but still it is a very reasonable match.

This almost concludes the discussion of the LNQA methodology. There is only one final aspect that I will address in the next section: Even if we have come to the belief that some causal generalization is reasonably accurate for the cases within its scope, how do we know *exactly* how strong the generalization is? And relatedly, how many cases do we actually need to consider before we are in a position to accept whatever strength comes out of our analysis? I will now consider these questions briefly, before I start to translate the NQA recipe into my own case study on paper.

4. Causal generalizations

The previous two sections established the process with which we arrive at a causal generalization. As I hope it is clear by now, this process operates through cases, and it is from those cases that the causal generalizations extract their validity. But I have not said anything about exactly how we get good, or *strong*, causal generalizations; neither have I said how many cases we might need before we can be confident of such a strength.

What do GH mean by 'strength' in the context of LNQA? The strength of a causal generalization in LNQA is a measure of its internal validity. In other words, it measures "the share of cases within the scope where the within-case analysis supports the causal model or theory" (148). This "share" is exactly defined as the number of cases that confirm the causal generalization over the total number of cases within the scope. Thus, the measure is normalized and can be used to measure the degree to which causal generalizations are supported by the cases within their scope.

Now, how do we stratify this measure? Is there a cut-off point where we can consider a generalization to be weak or strong? This cut-off point is, to some extent, arbitrary.

Acknowledging this, GH set it at 0.75, citing a number of institutions and research that also have a similar cut-off point at which they consider claims to be sufficiently general. Examples of this include the US Department of Labor, which according to the 4/5 rule, "for any protected group which is less than four-fifths (or 80 percent) of the rate for the non-protected group is taken as prima facie evidence of adverse impact", or QCA analysis, which uses a 0.75-0.8 cut-off as a "critical consistency" point (151). When the strength is 0.95-1, these are called "iron-laws" (151), a term derived from balance-of-power theorists (Levy and Thompson 2011).

What about strengths lower than 0.75? Here they take inspiration from Cartwright (no citation provided), who used Mill's word "tendency" to express "causal relationships which are nontrivial, but not particularly strong either" (151). They assign the character of tendencies to causal generalizations that fall within the 0.55 to 0.75 range. These tendencies are strongly related to "Straw-in-the-wind" tests in the process tracing literature, because finding a generalization of this strength is equivalent to slightly raising our belief in the generalization. Finally, generalizations below 0.55 are considered not-strong. The "strength curve" associated with these cut-off points is represented in Figure 9.

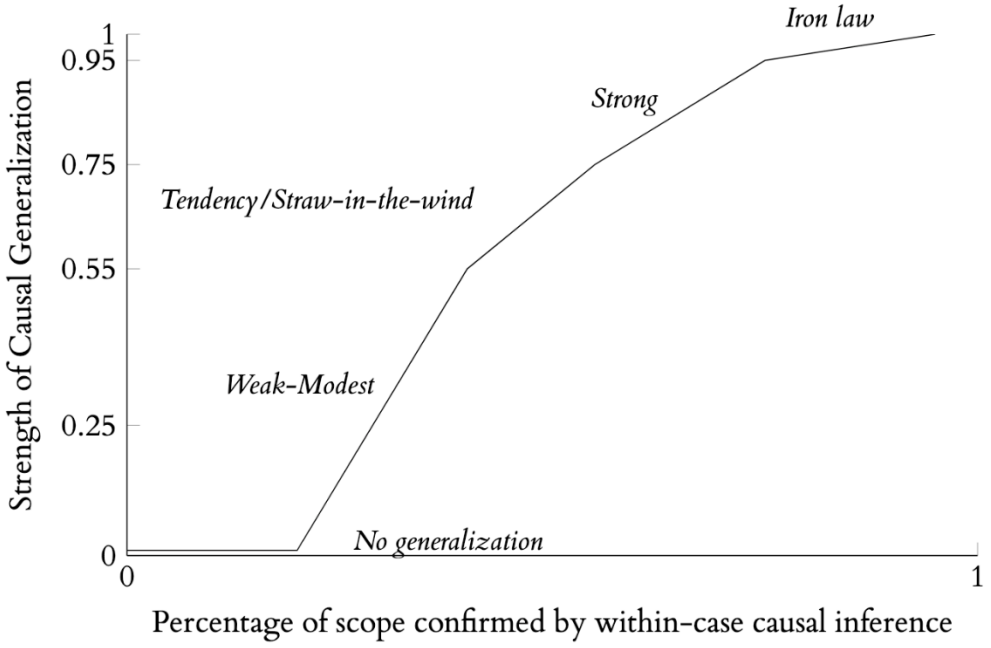

Figure 9. *Denominations of the strength of causal generalizations depending on percentage of confirmatory cases (149). [Now, I should be forthcoming and say that I am left quite confused by this figure. Since the strength of a generalization is exactly the percentage of confirmatory cases, this graph should be a straight line. Unfortunately I have not seen in the text any further explanation of the graph that would clarify what is going on.)*

The notion of the "strength" associated with a causal generalization, as I said, measures the internal validity of a causal generalization. However, the LNQA method on its own "by no means resolve problems of external validity", understanding external validity as

"whether the hypothesis extends beyond the stipulated scope" (157). GH make it very clear that it is "not incumbent" on LNQA to "establish the external validity of its claims" (158), and in some way that is reasonable: The point of LNQA is precisely to establish clear scope conditions that can set the cases on which the generalization applies—in doing so, it sets a clear boundary between the internal and external validity of the claim and focuses on robustly guaranteeing the former.

However, GH also point out that partly for that very reason questions about external validity tend to come up: If we have establish clearly that some causal hypothesis strongly applies in some domain, what about other domains that might be relevantly similar? While LNQA does not have an in-built toolkit for addressing questions of external validity, they suggest that the M-STOUT could be a helpful strategy. I will only briefly describe M-STOUT here, but the reader can find the more extensive discussion on pages 158-163.

Developed by Findley and colleagues (2021), M-STOUT is an acronym that stands for a number of "dimensions" along which we might want to study the external validity of a claim. S stands for "Setting", meaning "conducting the same analysis in a different setting in which features of that setting could affect the X – Y relationship" (159). In the context of (say) comparative politics, changing the country is equivalent to changing the setting. Imagine that Boucoyannis claim were restricted to England. Comparing the formation of parliamentary courts in France, then, would be like testing the external validity of the claim in a different setting.

T stands for "Treatment", and means changing the X to "values … that are outside of the scope of the study". Conversely, O stands for "Outcome" and refers to changing the Y in the generalization. (159)

U stands for "Unit", and although related to the Outcome dimension, is different in emphasizing changes in the unit of the outcome. For example, Mazur analyzed the Syrian civil war in (2021), centering the unit of analysis in the ethnic groups of the country. At one point, he considers whether his analysis might also apply to "other ethnically-dominated authoritarian regimes of a particular sort, i.e., units of the same sort elsewhere", as GH put it, stating that "The mechanisms examined here [ie., those related to state-society networks] are most likely to operate in cases where challenge originates domestically and during the period before a polity enters into prolonged intrastate war" (2021, 24, cited by GH, 161).

The second T stands for "Time", usually the future, seeing whether the analysis that has been performed is likely to stand at some later point in time. GH cite some studies (see

162) where scholars of conflict between dominant and rising powers consider what will occur to current US-China relations in the future.

Finally, the M stands for "Mechanism", a crucial component for M-STOUT (hence why it is singled out), and which GH consider the main reason of affinity with LNQA. "To some extent", they say, "external validity in LNQA always revolve around the question of whether a postulated causal mechanism works in the same way in different settings, with different levels of the treatment and so on" (162).

Now we know how LNQA deals with both internal and external validity. There is only one more question that we need to address before concluding this section: How many cases do we need to consider before we can feel confident in the strength of a causal generalization? Notice how this is a different question than how strong (i.e. internally valid) a generalization is. One can see this by wondering whether a 4/5 generalization is perhaps less robust than a 80/100 generalization. There is, then, a general question about the size of the population of cases that we need to consider. Is there some sort of heuristic cut-off that we might be able to use?

To answer this question, GH use Dion's (1998) work on sample sizes when doing comparative case studies. Dion approached the question from a Bayesian perspective in terms of necessary conditions: that is, what are the number of cases required such that $P(X = 1 \mid Y = 1) = 1$. He produced a table showing the possible combinations. (Figure 10)

		Alternative Hypothesis								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Prior Probability	0.1	3	4	5	6	8	10	15	23	48
	0.2	2	3	4	5	7	9	12	19	41
	0.3	2	3	4	5	6	8	11	17	36
	0.4	2	3	3	4	5	7	10	15	31
	0.5	2	2	3	4	5	6	8	13	27
	0.6	2	2	3	3	4	5	7	11	24
	0.7	1	2	2	3	3	4	6	9	19
	0.8	1	1	2	2	3	3	5	7	14
	0.9	1	1	1	1	1	2	2	3	7

Figure 10. *A Bayesian analysis of N cases required for a 1.0 generalization (155, based on Dion (1998)). GH wants us to draw our attention to the cell where $P = 0.95$. Assuming that both $P(X) = 0.5$ and $P(Y) = 0.5$ (uniform priors in standard Bayesian epistemology), this occurs using only $N = 5$ cases, a stunningly low result.*

The "striking" result from Dion's approach is that, if we want to be 0.95 confident that the generalization has a strength of 1 (i.e., that it is an "iron law" necessary condition), and

assuming nothing about the priors and thus attributing them a 0.5 probability, the number of cases required are only 5.

This is—I agree—a striking result that suggests (to me, at least) that LNQA is a more practicable methodology than may be suggested by the allusion to "Large" N. Just to reinforce the result, GH mention another study by Braumoeller and Goertz (2000) which arrived a similar results (specifically 5 to 6 N) from a non-Bayesian approach. I conclude with some words by GH that further reinforce these ideas about the sample:

> The LNQA answer to [the] question about "how many cases do I need?" is that medium-N designs can produce strong causal generalizations. This result is made even stronger by the fact that Dion's Bayesian logic is just about regularities without the accompanying LNQA requirement for case studies and within-case causal inference. If we add the results of extensive within-case analysis into the mix it makes these calculations even more compelling. (157)

Part II: LNQA in practice

In this Part, I will run a translation exercise and consider how an LNQA methodology would approach producing a causal generalization I am interested in. To do that, I will follow the same structure I did in the previous sections. First, I will consider what regularity I want to explain and might be worth tackling in a Large-N qualitative fashion. A crucial task at this point is to delimit exactly what the regularity refers to, and establishing scope conditions will be important. After that, I will start considering a causal story that might explain the regularity. To do this, I will present a causal diagram and briefly run a contained within-case analysis for a single case out of the entire scope. I will not consider issues about strength or external validity since I will only be assessing a single case, but towards the end I will suggest ways in which the study could be continued and expanded.

5. If paper, then… what?

In the introduction, I mentioned that the historical episode I would consider has to do with material culture in the history of Islamicate[3] science. By far, one of the material innovations that most profoundly impacted Islamicate science was the development of

[33] Quick semantic note. The term "Islamicate" was coined by Hodgson (1974, 59) to refer "not directly to the religion, Islam, itself, but to the social and cultural complex historically associated with Islam and the Muslims, both among Muslims themselves and even when found among non-Muslims". The broad phenomenon of Islamicate science, then, should not be understood as an inherently religious enterprise but rather a kind of practice that arose within a culture where Islam was the predominant religion. This distinction is important because many historical figures of "Islamic" science were not Muslims, such as Avicenna, and thus had much broader vocation than merely serving the interests of one religion.

rag paper. Paper was not a new technology, originally having been invented in China likely at least by the 2nd century CE and produced from the inner bark of the mulberry tree. The kind of paper that the Arabs exploited and exported, however, sourced the cellulose not from the bark of the mulberry tree, but rather from disposable rags like fish nets or old clothes. This seemingly small change in the manufacturing process had implications that reverberated all throughout Medieval Eurasia, as it massively lowered the economic cost of manuscript production.

As it happens, what is traditionally called the "Golden Age of Islam" (see e.g. Saliba 1995) sees its beginnings approximately at the same time as the introduction of paper in the Abbasid administration, during the Caliphate of Hārūn al-Rashīd (d. 809). Theories abound as to why the technology was implemented, which I will not get into. What matters is that, as soon as paper starts being employed in the administration, its use spread like wildfire across the Islamicate world. By the 11th century, for example, the Egyptian papyrus industry had all but imploded under the weight of the demand for cheap and durable paper.

The historiography of Islamic science has traditionally linked the introduction of paper to the cultural explosion of the period. This tradition extends all the way to Ibn Khaldūn (d. 1406) and reaches to the present day to the most recent synthesis on the impact of paper in the Islamic world by Jonathan Bloom (2001). The underlying idea behind this scholarship is that the introduction of paper somehow affected the way that information was produced and communicated, thus affecting culture. Just to give a flavor of the kind of narratives that one can find in the literature, I will cite in length one of the starkest examples by Austrian Ortientalist Alfred von Kremer (d. 1889):

> From a cultural and historical point of view the reduction in the cost of writing material, which went hand in hand with the production of paper, was of great importance. Books on parchment or papyrus were so expensive that they were available to very few. By the production of a cheap writing material, and its supply to markets both east and west, the Arabs made learning accessible to all. It ceased to be the privilege of only one class, initiating that blossoming of mental activity which burst the chains of fanaticism, superstition and despotism. So started a new era of civilization. The one we live in now. (1920, 208)

Kremer's idea is that a cheap writing support like paper leads to a democratization of access to knowledge, via in addition a wide book market that spanned from the Eastern Atlantic all the way to the Indian ocean. Now historians agree, of course, that the causal link is more complex—knowledge could not have been democratized, for example, if a

wider public had not learned the Arabic language required to read and understand the Quran. But the core idea that the introduction of paper somehow altered the conditions of cultural production, thus perhaps the products of such a culture, persists.

Unfortunately, there is little if any scholarship on how exactly such a change was produced. Access to a cheap writing material does not automatically change the products of a culture. We need to hypothesize some mechanism that can account for such a change. In this way, we have a basic association—change in writing medium is associated with a change in culture—but no robust causal account that can more or less confidently explain such an association. In what follows, I will sketch how LNQA could help us tackle this explanatory project.

To do this, first let's formulate the regularity that we want to explain. The first thing to say here is that I am interested not so much in Islamicate culture in general, but rather in their scientific production. The question is then how to link the introduction of paper to specifically scientific productivity. To frame the discussion and the choice of cases, I will consider a series of nested regularities. I will start with the most general expression of the phenomenon and progressively narrow it down through strategies that GH cover in their textbook. We could express a first-pass regularity as:

R1. *When paper is introduced into scientific workflows (X), there is a development in science (Y).*

This is the most general regularity I can design for the cases comprehended in the Medieval Islamicate world. At least from a superficial point of view, it seems to match the basic association I mentioned earlier, which related the introduction of paper to the scientific developments of the Classical Islamicate period. The problem with this claim is that it is too broad and unmanageable. Part Y of the claim is so big that it will include several disconfirmatory cases—at the end of the day, paper becomes the main writing support in the Islamicate world by the 11th century (with relevant exceptions, as we will see), and it is to be expected that science did not develop absolutely everywhere. Indeed, the notion of a scientific "development" is hard to pin down. Similarly, X is too wide-scoped not to include potentially infirming cases. Nevertheless, I emphasize R1 as an important motivational claim.

Let's start by nailing down the scope of X. Here I argue that we need to narrow the scope of the claim both spatially and temporally (see Garry Goertz and Haggard 2025, 28 for a discussion on geographic strategies). Introduced, when? And in what locales? Since my motivation is to provide some explanation of how paper contributed (if indeed it did) to the

scientific developments of the Islamicate Middle Ages, this provides some opportunity for narrowing.

The region of the world where Islam had some sort of cultural hegemony, even by the 9th century, is enormous, encompassing territories from the Atlantic to the Indian Oceans. (Figure 11) As we move towards the High and later Middle Ages, these borders both expand and contract. In 1206, Muhammad Ghori takes over the North India planes and founds the Mamluk Sultanate, consolidating the political presence of Islam in the Indian subcontinent. By the end of the century, several Islamic polities had also sprung up across the Northern parts of Sumatra, extending the reach of the religion and its influence into Southeast Asia. However, Islam's dominion in Europe was also starting to fade. Sicily had already become a politically Christian-aligned territory by the end of the 12th century, and at that same period the Islamic powers had started irreversibly retreating from key locales of the Iberian Peninsula, to be finally "reconquered" in 1492.

Figure 11. *Territorial bounds of the Abbasid caliphate and satellite states in the 9th century (Afsaruddin, 2025). Only the Westernmost region of the Maghreb (effectively al-Andalus—"Spain" in the figure—and the Idrīsids in current day Morocco) had political independence of the Abbasids. Culturally, of course, the center of Baghdad exercised an enormous pull across the Islamicate world, including the Maghreb.*

All of this makes choosing a geographical and temporal bound to the claim a very complex issue, and I have no choice but to introduce a level of arbitrariness in how I cut up the regions and period. Nevertheless, I think the period that goes from the 9th to the 14th centuries casts wide enough of a net to be sensitive to the changes in scientific practice that are caused by paper, without too much territorial flux that renders the problem overwhelmingly complex. By the 9th century paper is already established as a writing

material in Baghdad and is rapidly taking over other areas of the Islamic world. Conversely, by the 13th century the pioneer regions in the use of the technology have all but ditched other writing materials, and regions like al-Andalus have fully integrated paper into their scientific production processes. Within these 4-5 centuries there are key changes in scientific writing practices that could be attributable to paper (for instance, several mathematical textbooks in the Islamic East are describing arithmetic computations by including all intermediate steps in the writing surface, see Berggren (2016)), and there are a number of regions with sufficient continuity that we can follow their evolution.

What would be the units of X that we want to consider? Here a dose of arbitrariness in the classification will also inevitably slip in. Clearly, an astronomer writing in al-Andalus should be considered a different case than a physician writing in Khurasan. But what about one writing in Fez and another in Palermo (*Balarm*)? Given that we have to slice a very big territory in some way, we may choose to pick political divisions as a starting point. But this has its limitations—It has been convincingly argued, for example, that there is a consistent astronomical tradition in the entire Western Islamicate world, including al-Andalus and the Maghreb (Samsó 2020).

To cut to the chase, I think there are at least five regions that, for different reasons, would require separate analysis: The Maghreb (including al-Andalus), Egypt, Syria, Iraq, and the Persianate region of Khurasan. The selection of these specific units responds to a number of selective interests in certain dimensions of Medieval Islamicate society. Fundamentally, I want to be selective to the cultural features that can have a bearing on regional differences in scientific practice or products. I am thus interested in regional differences in linguistic and religious heritage, as well as the predominance of intellectual traditions. Khurasan and Syria, for example, are the two main bearers of the Syriac and Persian groups within Medieval Islamdom. These two groups, in turn, provided the sources for the translations of the 9th century and their dynamic characterized the scientific developments of the “Golden Age”. (Saliba 2007) In other cases, such as Iraq and Egypt the political autonomy merits separate treatment. The Maghreb, on the other hand, deserves separate treatment because of its own peculiarities in the way that science developed in the region (plenty of detail for this claim, especially in what touches the

material infrastructure of writing, can be found in the dedicated chapter on Medieval Iberia).

An LNQA researcher might consider the evolution of scientific writing in these five regions, which given Dion's results (see Figure 10) might allow one to be confident in the resulting strength of the generalization. This gives the regularity:

R2. *When paper is introduced into scientific workflows in the Islamicate world from the 9th century and into the 14th (X), there is a development in science (Y).*

Now, what about the effect (Y)? I suggest we nail down part Y of R2 via conceptual engineering (see GH, 25). I argue that what we need in the effect (Y) is not so much a change in the *products* that come out of the scientific endeavor, but instead a change in the *practices* which are more directly affected by changes in the material culture of production. Note that this is not a trivial statement. In places where papyrus or parchment were already in use (and this was the case in the entirety of the Islamicate and European regions) the introduction of paper does not imply automatically a change in practice: The crucial tandem of pen and some pliable, lightweight surface does not change. However, there *was* a change in practice. To account for such a change, we will need some further contributing causes, which I will consider when I start hypothesizing explanations for the regularity. We could, then, narrow the regularity with:

R3. *When paper is introduced into scientific workflows in the Islamicate world from the 9th century and into the 14th (X), there is a change in scientific practice (Y).*

But the effect (Y) is still too wide-scope. Scientific practice includes plenty of practices that most likely will not be affected by a change in the writing materials. One need only think about the fact that a lot of science in the Middle Ages was made in an oral and aural context, via dialogical reasoning, memorization, and communal recitations. Famously, the classification of animals up to at least the 15th century was heavily influenced by mnemonic techniques (see Carruthers 1992, 161), so it is reasonable to assume for instance that a change in the writing infrastructure did not change animal classification—at least until new methods and ideas in the discipline arise in later centuries. I propose, then, to further narrow down the effect in the following way:

R4. *When paper is introduced into scientific workflows in the Islamicate world from the 9th century and into the 14th (X), there is an expansion of astronomical intermediary documentation (Y).*

By "intermediary documentation" I mean all form of writing that, while contributing to a final manuscript or book, does not end up in the final version. Two crucial examples that are my main focus in the context of paper and science are the annotation of observations

and the use of arithmetic techniques for the processing of quantitative data. Consider for a moment a finalized version of a *zīj*, astronomical handbooks popular in the Islamicate Middle Ages that contained astronomical data for the use of all kinds of professions, from astrologers to navigators. These *zījes* did not contain the computations that allowed their authors to arrive at the quantities displayed in a *zīj*. These were steps in the process of production that had to be done in a different writing support, perhaps a support more suited for the kind of ephemeral work that goes into such intermediate stages. R4 suggests the possibility that paper was instrumental in that stage of production. Finally, I consider astronomical due to the fact that some of the practices I will consider, like quantitative work, most directly benefit sciences that were quantitative at the time. Even though I will restrict my analysis and claims to astronomy, I should say that I see no reason in principle why the results I arrive at here might not extend to other disciplines in the period.

R4 is the final form of the regularity I argue would be amenable to examining with LNQA. All regularities I have suggested (R1-R4) are expressed as X regularities. What this means is that they portray the causal contribution of paper as sufficient for the expansion of intermediary documentation in the Islamic world. But why use an X generalization to understand the relationship between paper and scientific practice, rather than a Y generality?

In this case, it is crucial to bear in mind the scope conditions of the claim that restrict the phenomenon we are interested in. We want to provide a (hopefully) generally valid mechanistic basis for the effect of paper on Medieval Islamicate science. That is, we want ultimately to construct a causal generalization that explains the mechanism by which paper influenced the development of Medieval Islamicate science. Ultimately, one can imagine many reasons why intermediary documentary practices may expand, both in the Medieval Islamicate world and elsewhere. The very transition from an oral and aural tradition in the Early Medieval Islamicate culture to a written one can lead to similar expansions in documentation. (Schoeler 2009) Similarly, in Renaissance Europe an increasing drive towards systematization played a major role in the production of notes. (Blair 2010)[4] The focus is not, then, what was necessary for the expansion of

[4] It is an interesting fact that in both examples paper most likely played a major role: In the Islamicate case, by providing a cheap medium on which to write; in the European case, by promoting the production of information that humanists felt the need to organize. To my point, though, this does not mean that paper is a *necessary* condition. It is very reasonable to suppose, for example, that papyrus may have had similar effects on the development of Ancient Mediterranean knowledge practices. And yet, I cannot skip the opportunity to note that in the context of the Ancient Mediterranean, papyrus was functionally a very paper-like technology: relatively durable and cheap, and very easily handled and transported. Perhaps what matters is the functional characteristics of the material technology in the social context in which it appears. More on this in the next section and especially in the chapters that follow.

documentation in the Medieval Islamicate period, but rather whether the introduction of paper in this specific context can explain the expansion of intermediary documentation we see in the historical record. In other words, we want to know if X is sufficient for such an expansion—hence why I opt for an X regularity.

Summing up, we want to follow traces of expansion of intermediary documentation in the Islamicate world from the 9th century to the 13th century, which includes at least the Maghreb, Egypt, Syria, Baghdad and Khurasan. In all cases, we need to check whether: 1. Paper is introduced into their scientific workflow (X), and 2. Astronomical intermediary documentation expands (Y). While a detailed account of each case is beyond the scope of this text, I will shortly provide some key pieces of information that can helps us determine the strength of the regularity. I will present the cases in order of arrival of papermaking.

Paper and papermaking reached Khurasan first, most likely by the fourth century as a result of the trade routes that connected East Asia to the Mediterranean. (J. M. Bloom 2005) And the 9th century, private observatories were already in operation that required the collection of observational documents from different parts of the region, for example in Nishapur. (Morelon 2019) Morelon makes a case for these private observatories in the context of an unprecedented expansion of observational practice, going beyond the record set by the work of Ptolemy: "It is therefore evident that, from the ninth century, the Arabic astronomers had at their disposal the results of a far greater number of recent observations than those available to Ptolemy when creating his work" (2019, 9).

From Khurasan, paper arrives to the Baghdadi Abbasid administration. The geographer Yāqūt al-Ḥamawī (d. 1229) tells us in his *Muʿjam al-Buldān* that the first papermill in Baghdad was set up during the caliphate of Harūn al-Rashīd, around 794-5 AD. (ibn-ʿAbdullāh al-Rūmī al-Ḥamawī 1861) This coheres with the account of ibn Khaldūn that I presented earlier. Once the Abbasid administration incorporates paper into their administration, a number of documentary practices experience expansion. Perhaps the most interesting one in this case is the growth of scientific translation, which starts with the caliphate of al-Maʾmūn in the 9th century AD (for details, see Gutas 2012).

Papermills existed in the Syrian region (Hama and Tripoli) since at least the ninth century. (J. Bloom 2001) This paper was widely used and exported, as testified by a 9th century manuscript in Greek on the *Doctrina Patrum*. (Vatican Library, Gr. 2200) Intermediary scientific documentation similarly expanded in the region. In 14th-century Damascus, the institutionalization of mosque timekeeping correlates with the production of large, working astronomical corpora like prayer-time tables and universal auxiliary tables. Al-Mizzī (d. 1341), for instance, compiled tables for timekeeping and prayer times for Damascus that survived as a compact set of tables in Cairo MM 62, which al-Khalīlī

expanded in his famous universal tables (best preserved copy of 1411 in MS Paris BnF ar. 2558). (King 2018)

The date of arrival of paper in Egypt is approximately the 9th century. Karabacek (2001) argued that the oldest paper in Egypt is a letter from 874, and to my knowledge there have been no older paper found since then. Egypt is also the region that presents the richest evidence regarding the uses of paper in the Medieval period due to the existence of the Cairo Geniza, which presents a very rich collection of scrap papers and all sorts of intermediary documentation—scientific and otherwise—from the period. (Goitein 2001) In my case study on the Iberian peninsula I consider two materials that reflect scrap astronomical calculations on paper: the JRL Series L 22 1 / 1 leaf, recto and the T-S K3.19 1r (for the digitized versions of these fragments, see the searchable database of the Princeton Geniza Project).

More broadly, the abundance of intermediary scraps in the Cairo Geniza testifies, I believe, to the broad existence of new practices of intermediary documentation across the Medieval Islamicate world. As has been amply discussed, we should thank survival of these scraps in the Cairo Geniza to the fact that, since they contained the word "God", the Jewish religion prohibits their destruction and thus were archived in a form of ritualistic burial. (Hoffman and Cole 2011) While this may bias the content of the fragments on some level, a good portion of them contain text written on resold paper—which of course did originally have that bias in its content. In this sense, the Cairo Geniza can be read as a *random sample* of the documentary practices of the Medieval Islamicate period—and a representative one at that, insofar as it includes writing from all of its regions. The abundance of intermediary texts in the Cairo Geniza, then, provides high-quality evidence about the broader writing practices of the period and its people.

Finally, we have the case of the Maghreb. I will not get into that in this section, since I consider that case in specific in the next section and the case study of the dissertation, where I explain in detail the expansion of intermediary documentation due to paper that occurred there. While I have not yet provided much reason to believe it, I consider the Maghreb as conforming to R4.

Figure 12 shows the cases sorted out in a two-way table. The reader can see that 5 out of the 5 (100%) cases conform to the regularity, making R4 very strong. As seemingly unlikely as this might seem, I should point out that this should not be that surprising. As I pointed out at the beginning of this section, the idea that the introduction of paper led to new documentary practices is taken as a given, and that in doing so it perhaps had an impact in the development of Medieval Islamicate science. The key question is not, then, whether paper contributed to science—like it did to so many other realms of culture—but

exactly *how* it did this. In other words, the main problem is the *mechanism* with which paper impacted culture. In the next section, I will address that particular question in the context of scientific practice.

	Paper is not introduced into scientific workflow (X=0)	Paper is introduced into scientific workflow (X=1)
Expansion of intermediary documentation (Y=1)	—	5
No expansion of intermediary documentation (Y=0)	—	0

Figure 12. *Two-way table for testing association R4. Since we are interested in testing an X regularity, we care about the cells where X is present (right column). There are no non-conforming cases to the regularity, rendering it very strong.*

6. From paper to intermediary documents

As I discussed earlier, for GH it is imperative in any LNQA to produce a causal diagram before any causal claim is to be taken seriously. This is because the diagram guides the search of evidence, looking into the presence or absence of the cause(s) and effect(s), as well as the mechanism proposed to operate in the cases. Once the diagram is produced and the causal hypothesis made clear, an LNQA analysis proceeds to look for conforming or nonconforming cases within the scope of the regularity.

I propose the causal diagram of regularity R4 in Figure 13. The two key components are the cause (X) and the effect (Y). Since I treat X as sufficient for Y, I consider a set of mechanisms (or mediators) that take us from X to Y, as well as some relevant moderators that I argue affect how the mechanisms bring about the effect.

The main idea, again, is that the introduction of paper into the scientific workflows of the 9th-14th c. Islamicate world led to an expansion of intermediary documentation. This causal diagram is meant to explain exactly how that came to pass. The temporal structure of the claim—that, given the introduction of paper in the 9th century, by the 14h century intermediary documentation expanded—sets the X as "Paper is introduced into the 9th c. Islamicate world", and the Y as "Expansion of intermediary documentation by the 14th c.". Following GH, I take causal diagrams to reflect a temporal direction, so whatever is between the cause and effects are events that happen at some time between the occurrence of the cause and the effect.

Let's start by focusing on the cause X. As I mentioned, I treat the introduction of paper as a sufficient cause of the expansion of intermediary documentation in the Medieval Islamicate world. However, the introduction of paper assumes that the material is actually

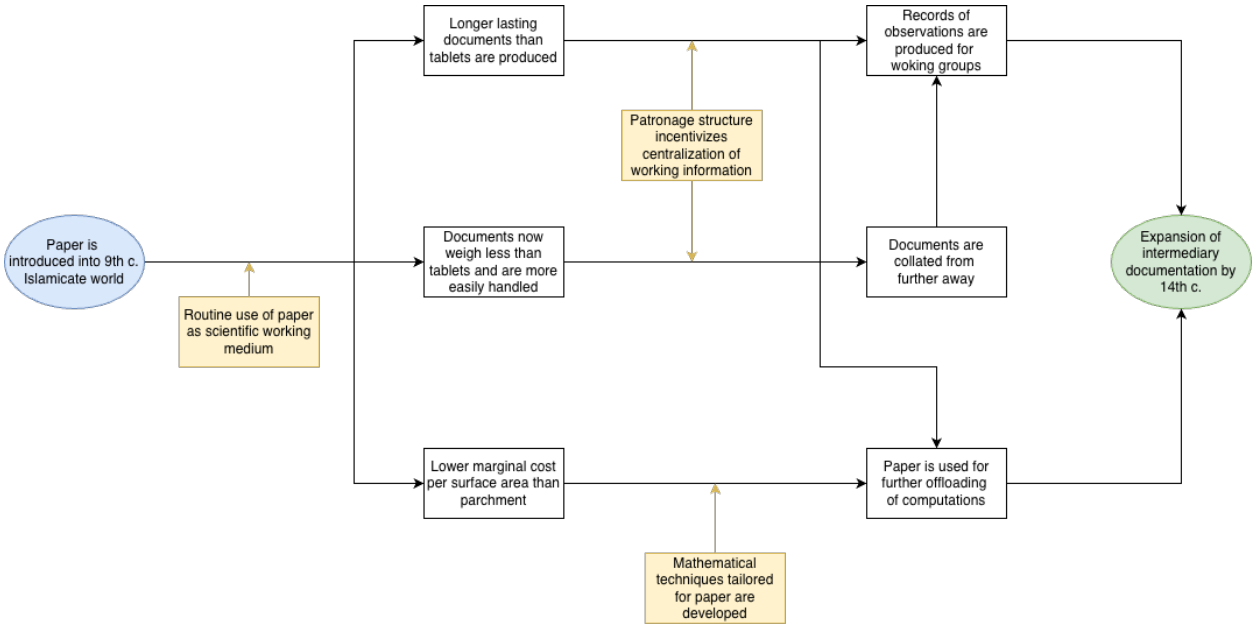

Figure 13. Causal diagram of regularity R4. In blue: Causes (X) that are treated as independent variables. In green: Effect (Y) that comes about from the cause and its associated mechanisms. In white: Mechanisms which bring about the effect. In yellow: Moderators that affect the extent to which the mechanisms will exert their intended effect. Causes or mechanisms are treated as jointly sufficient for a particular outcome whenever arrows converge on the outcome.

being used in the process of scientific production. I thus include a moderator for the cause X. Once paper is routinely adopted for scientific work, then we can start assuming that some of the properties of paper will allow practitioners to adopt some key practices. It is part of the within-case role to establish that paper was not only present, but actually used in the context of production. This is not trivial, because paper has indeed many uses, from wrapping commodities to covering graphic representations on the walls of a mosque. (CITE) But paper being used is, of course, dependent on the presence of paper itself, hence why I treat it as a moderator rather than an independent cause.

In my case study on paper, I highlight three key properties of paper that played a role in the new writing practices of the period. Paper is, compared to the other available writing materials in each practice where they were deployed, more durable and manipulable than wax and dust boards, and especially cheaper than parchment. As a result, longer lasting documents can be produced, and those documents weight less than those written on tablets and are more easily handled. Crucially, the introduction of paper lowers the marginal cost per surface area of written document. This is important, because it strongly incentivizes writing that until then was not done due to financial constraints.

The crux of my hypothesis is the idea that each of these properties of paper led to specific writing practices in the context of the scientific workflow. Let's go one by one. If paper allowed for the production of longer lasting documents, this incentivized the

production of records of observations in the context of working groups of practitioners. There are, of course, advantages to having other people make some observations of a stellar trajectory—say, the trajectory of the Sun—but these advantages cannot be reaped if practitioners don't have a collective archive to refer to. Having paper available makes storing these documents more feasible, thus incentivizing their production. This applies not only to their durability, but also the fact that paper can be stacked and archived much more easily than tablets—this is assuming that hoarding tablets in the thousands would make sense for any purpose, something unheard of in the historical record.

If documents are more lightweight, that also makes it easier to transport them across longer distances. This being the case, one should expect that paper documents not only travel longer distances than tablets (or even parchment, given that they are a more valuable resource), but that they are more concentrated within a single locale wherever working groups operate. This should also incentivize the production of more documents, given that they can be shipped for reference and use. Since I am interested in scientific production I focus on the production of observations, however these ideas may apply to other kinds of documents.

In the diagram, I include a moderator that applies to both mechanisms of durability and manipulability. The moderator introduces the factor of patronage: "Patronage structure incentivizes centralization of working information". Now, in some ways, patronage has been a constant moderating factor across the history of culture, including today, and it has taken different forms in different contexts. Here I am interested in underlining one factor it introduced in scientific practice in the Medieval world—not only Islamicate, but also European—namely the fact that it centralizes scientific activity within the locale of the patron. The level of centralization can vary, of course. While in some cases like Alfonso X's *scriptorium* all scientists working for the monarch physically worked in designated premises (Román de la Higuera (d. 1611), for instance, speaks in his *Historia eclesiástica de la imperial ciudad de Toledo* about the old alcazar as the site for the observations that led to the Alfonsine table (Mss. 1289, Book 22, Chapter 12)), in other cases we may have different practitioners working in different locations with overlapping projects. That seems to have been the case for Said al-Andalusi's patronage (see next chapter for details).

In both cases, however, the locale of the patron becomes the central node of information exchange, where documents are brought and information is put together. To the extent that such a patronage structure exists, it will incentivize such physical collation of documents. However, it is important for the purposes of our causal model that the collation of documents is not the only mechanism leading to an expansion of intermediary documentation. In this sense, patronage moderated an important contributor to the

presence of the effect, but the effect could still be present were the patronage structure to dissolve—albeit, perhaps with lesser strength.

The third property of paper that is key in the expansion of intermediary documentation is its relatively lower marginal cost in comparison to parchment. As we saw with the case of von Kremer, this has traditionally carried the weight of most historiography on the impact of paper on Islamicate science. I take on some of these arguments, but I make them concrete by arguing that a cheap writing support motivates its use for computational offloading, thus generating more intermediary documentation. Note that the durability of paper also contributes to this outcome. The space of a tablet greatly diminishes how much of a computation one can include in the writing surface before what is written has to make way for further computation. A cheap, durable material like paper incentivizes offloading previous computational steps that the computer can come back and refer to. I suggest this might be especially relevant as the computations become arduous and complex.

In the diagram, I add a moderator that mediates the transition from lower marginal cost to the use of paper in offloading. This moderator introduces the condition that, beyond paper being introduced, new mathematical techniques have to be developed in order to use them for offloading. The key idea here—to be much more thoroughly explained in the case study chapter—is that, previous to paper, computations were done on smaller and less replaceable materials (such as wax and dust boards) that incentivized an arithmetic of erasure, where in order to carry out complex or laborious operations the computer has to erase previous steps to make way for the new ones. As paper is introduced around the 9th century, one can trace the progressive offloading of intermediary computational steps in the manuals of elementary arithmetic of the Islamic world. Of course, all of this assumes that practitioners will develop the new arithmetic techniques to make use of the extra room—hence the moderator. To the extent that the new techniques are developed, paper's lower marginal cost will be able to exert its action in the expansion of computational documentation.

Overall, the introduction of paper into scientific workflows should promote the intermediary documentations of at least two kinds of documents: Observation logs and computational work. They could be more—I have said nothing about drafts, for example—but at least those two are basic in astronomical work, and would thus lead to a more general expansion of intermediary documentation in that regime. Although I do not want to make a claim about progress in Medieval astronomy, it should also be noted that these two kinds of documents buttress an expansion in the infrastructure of data sourcing (via observational work) and data processing (via computational techniques). This is relevant at

least to better understand the role of new writing infrastructures in the context of Medieval developments in astronomy.

Many authors (see e.g. McCluskey (2000) for the European case, Saliba (1987) for the Islamic case) have suggested that the Middle Ages witnessed a profound systematization of quantitative astronomy. Central to this systematization is the interplay between tabular parameters that are constantly being checked against the astronomer's own observations. Particularly the Islamicate practitioners engaged in a long, far-reaching debate about the compatibility of the Ptolemaic tradition with observation, buttressed by new computational methods like the new trigonometric functions inherited from Indian mathematics. Medieval astronomy (especially as the Islamicate tradition conceived of it), then, introduced the connection between observation and computational methods as two sources of information that are meant to cohere for a realist depiction of the heavenly bodies. In this context, the expansion of observational and computational documentation could be considered at least as consistent with the new impetus to quantitative practice provided by the Medieval, and especially in my period of interest Medieval Islamicate, astronomers.

Let me make a caveat before we consider the within-case evidence. The diagram from Figure 13 should not be considered "complete", in the sense of the "completeness" criterion for causal diagrams (see p. 11). This is mostly for practical reasons, although it is also due to reasons of focus. In the same way that Eisenstein wanted to isolate print's "special effects" (Eisenstein 1980, xvi), I am also interested in what special effects paper had in the Medieval scientific process. It is for this reason, for instance, that the development of new arithmetic techniques for paper is a moderator and not a mechanism on its own. Had my focus been different—say, if I wanted to understand the mechanisms that led to the mathematical inventions of the period—the new arithmetic techniques might have been a step in the causal process and the introduction of paper a support factor[5] for the creation of those techniques. To focus on paper's "special effects" requires putting the mathematical inventions as moderating its effect, but in other contexts those might have taken the role of a mechanism proper.

This lays out the general causal explanation for the regularity R4. Now the question is: Do the cases within the scope of R4 support the causal argument I have provided? Here I will not run a proper analysis of each case. However I will illustrate how the within-case causal inference would go for one case: the Maghreb from the 11th to the 13th centuries. The reason for choosing this particular case, which I will elaborate on further in later writing,

[5] Here I am using Cartwright et al.'s (2006) terminology. A support factor is for them a condition "needed if the cause is to produce its effect" (7). In our case, the introduction of paper is a condition needed for the development of arithmetic techniques that use more writing material.

can be cashed out in terms of smoking-gun tests for process tracing. Let me briefly say why.

When one looks at a map that portrays the diffusion of papermaking across Eurasia, it is easy to get the impression that paper spread with more or less uniform speed and penetration across the continent. (Figure 14) However, the diffusion of paper leads to areas of different levels of penetration, with the Maghreb (including al-Andalus) using a more diverse set of materials. In al-Andalus, paper was used alongside both parchment and tablets. Parchment, specifically, had enjoyed a strong supporting industry in the peninsula since before the Islamic conquests, but it also enjoyed a religious aura that rendered it more appropriate for the production of Bibles and Qurans. (On the parchment-making industry, see Vallvé Bermejo (1980), on the use of parchment in the Maghreb, see Déroche et al. (2015).) One could thus imagine that in the Maghreb the impact of paper would have been somewhat constrained by the presence of those other writing materials, and if we can see the mechanism from Figure 13 in action there, we might have more reason to expect it working in places like Baghdad or Khurasan.

Figure 14. Spread of paper across Eurasia from Central China around 2nd century BCE. (Bloom, 2017) By the 9th century paper is consolidated in the area of Baghdad and the Eastern Mediterranean, and it reaches Europe through Iberia by the 11th century. Diffusion of the technology is not equivalent to total takeover, however, and areas such as the Maghreb and Europe used parchment and other writing materials well into the 1600s.

In this sense, the X variable (paper is introduced by the 9th century) is less likely to occur in the peninsula. But the Y variable (that intermediary documentation expands as a result) is also comparatively less likely. This is because, generally put, the introduction of a writing material depends in no small part on whether there is a demand for it in the context of some practice. In the context of the peninsula, this means that writing practices were considerably attached to their specific writing infrastructure, which was mostly based on parchment and dust board at least by the Early Medieval Period. When paper shows up, it would have to battle an ideology built on the religious predominance of parchment, which

not only builds a bias against the use of other materials, but might also offer relatively more competitive prices that make the economic calculus fare slightly better for parchment than in other locales.[6]

In this way, the Iberian peninsula functions as a *smoking-gun test* from process tracing (see Bennett 2004, 210; originally from Van Evera 1997). As I already discussed, these are tests designed to test sufficient conditions for a hypothesis. In our case, the hypothesis would be that R4 is the case because of the causal story I described in Figure 13. The Maghreb functions as a sufficient condition because, given the fact that it is one of the locales where paper was less generally used, it can set a baseline for the level of impact that paper had in scientific production. If the Maghreb used paper, then in places like Baghdad or Egypt where paper took over all aspects of production, we should expect the impact to be even more thorough.[7] In this sense, the presence of paper use in Maghreb would be sufficient to show that paper was used everywhere else within the scope of our claim.

It should be said that it is not a decisive smoking-gun test, in the sense that it does not automatically demonstrate the hypothesis. One should still consider the other cases in detail before coming to general conclusions, especially in historiographical contexts, where other aspects of culture may substantially alter the way that actors relate to their material environment, including material culture. Nevertheless, I think it can safely be said that the Maghreb is the *least* likely case to conform to R4, and thus it raises the likelihood of the hypothesis the highest, were it to indeed conform to it.

So, does it conform to it? I argue that the answer is yes. While I do not have the space to explain why in detail (see separate case study for discussion), here I will briefly touch on the mediator (or mechanism) nodes of the causal diagram (Figure 13), and argue for their presence in the case of the Maghreb, with a focus on the Iberian peninsula, during the 11th to 13th centuries.

---

[6] It should be said that we do not have (at least to my knowledge) anything like a full quantitative account of the prices of parchment in the peninsula in the period we are interested in. Nevertheless we do know from sources such as the "Calendar of Cordoba" (*kitāb al-anwā'*) that parchment remained an active sector of manufacture for the Umayyad Caliphate. (Pellat and Dozy 1961) This makes sense in the context of the leather industry active in the peninsula, with products such as "Cordoban leather" or "guadameci" achieving great reputation and being widely exported, the first to Europe and the second to the rest of the Islamicate world (see Constable (1994, 192)).

[7] And in fact we know that the adoption of paper was much swifter in other places where data is available. We know this by the fact that the price of books went down very significantly during the Medieval period in places like Egypt. Shatzmiller's team has collected data on a number of aspects of the Medieval Islamic economy, including the price of books, that illustrate this point. See Shatzmiller (2018) for a synthesis on the economics of the adoption of paper in the medieval Middle East.

The exact date of introduction of paper into the peninsula, while debated, has an upper bound of the late 10th to early 11th century. The oldest remaining paper document from the peninsula (Ms. 6º Breviarium et Missale Mozarabicum) dates from at least the late 10th century (Valls i Subirà 1978, 102), and Levi-Provençal found a paper manuscript of the *mukhtaṣar*, or abbreviation, of al-Ājurrī's *Book on Sharia* dated from 970 that was part of al-Hakam II's library (1934). This is significant because courts are *less* likely to use paper, instead preferring parchment due to its prestige. This pattern repeats itself through Alfonso X's court, who was "caught" using paper for official documents and having those documents being sent back for rewriting on parchment (or *cuero*). (García Díaz and Montalbán Jiménez 2005) Given this trend, it becomes overwhelmingly likely that paper was commonly used in more routine scribal tasks, including scientific ones, outside of courtly environments where cost was much less of a constraint.

The three main material properties of paper that I argue causally contributed to the expansion of intermediary documentation (durability, manipulability, and low cost) were gathered from primary sources from the broader Islamicate world. These sources range from al-Thaʿālibī (d. 1038 CE) to Ibn Khaldūn (d. 1406). Al-Thaʿālibī, for example, describes paper as a better-looking, more "supple" material that is "more easily *handled* and is more convenient to write on" (1968, 140, my emphasis), signaling an appreciation for paper's "manipulability", as I call it. This makes paper not only more easily handled for short-term manipulation, but also long-term transportation across longer distances. Ibn Khaldūn, on the other hand, provided the oldest account of the introduction of paper into the Abbasid administration, alleging the change of policy to the shortage of documentation that arose from a reliance of parchment. (2015, 328)

Now, it should be acknowledged that these sources are not exactly located in the Iberian peninsula. Al-Thaʿālibī lived in Nishapur (modern day Iran) throughout his life, and Ibn Khaldūn, although originally from Seville, wrote most of his work in the Maghreb. This spread of the sources with respect to the material properties of paper has, I think, an inevitable weakness and a strength. On the one hand, the hypermobility (and disappearance) of Medieval Islamic documents means that so far we haven't studied much of the extant sources of the period. (Paul 2015) On the other hand, the sources we have about the Islamic use of paper do speak to a broader, shared experience with the product. Taking these sources together, and considering the fact that the basic product was the same in all cases (rag paper), allows us to infer that those characteristics were shared across the actors of the period.

The durability and manipulability of paper play a key role in the suggestion that working groups produced more records of their observations. Some of the most representative

astronomical projects undertaken in the peninsula from the 11th to the 13th centuries were, in fact, organized by working groups under the direction of a patron. Although we do not have sources that speak about the exact functioning of such groups, we get some inklings from tangential references by a variety of authors. Isaac Israeli ben Joseph wrote in his *Yesod* Olam (*The Foundation of the World*) in 1380 of a group who

> lived in the city of Toledo and other lands of Spain [and] devoted themselves to studying and experimenting through the observation of the heavenly signs. … Ibn Ṣāʿid [al-Andalusī, d. 1070] … loved science and whoever devoted themselves to it … and he offered what he possessed to his collaborators, providing for them and giving them stipends. … This group installed in the city of Toledo various precision instruments to observe and regulate necessary for the verification both of the position of the planets and their positions in the Zodiac, as well as their speed in each era and the duration of the solar year and the lunar month. (*Yesod Olam*, book IV, ch. 7, translated from Spanish version by Llavero Ruiz, 2000)

In this text, Isaac Israeli ben Joseph describes a scientific structure of patronage established in the Iberian city of Toledo. But he mentions in passing that this group also resided in "other lands of Spain", suggesting that other members might do their scientific work elsewhere, as is corroborated by the writings of Ṣāʿid al-Andalusī himself when he provides the genealogy of Andalusi astronomers from Maslama to his time. To do this kind of work, documents needed to travel into Toledo from other parts of the peninsula, where presumably the Toledan members would benefit from their "observation[s] of the heavenly signs".

This is corroborated in the writings of another member, commonly known as al-Zarqālī or in Latinized *Azarchel*, who talks about a group composed of "trustworthy individuals" that he relied on for the production of his *Comprehensive Work on the Solar Year*. But he adds another interesting fact—namely that the observations were collated "for the span of 25 years" (Cited in Millás Vallicrosa 1943, 279). This requires a coordinated storage of intermediary documentation for a considerable amount of time, given that 25 years sets a lower bound to how long these documents had to be stored.

The storage of these documents could only be in parchment or paper, at least on the assumption that tablets were not used for this purpose. Not only does this make common sense (I invite the reader to imagine a stack of 1000 folia next to a stack of 1000 wooden tablets), but it matches the fact that we have no records of an archive of tablets during the Medieval Islamic period. Conversely, we have plentiful references to paper archives, not

only in al-Hakam II's library but also in other chanceries around the Islamicate world. (Hirschler 2016)

Overall, we get the sense that astronomical practice in 11th century Iberia benefited from paper in what concerns intermediary documentation. Observations had not only to be produced, but they had to be collated to the patron's center in Toledo from further away in the peninsula, and stored in paper for further use and reference. This is an entirely different *modus operandi* than the practices characteristic of previous astronomical work in the peninsula.

The last part of the within-case inference concerns the expansion of computational writing due to the introduction of paper. There are several avenues that could be pursued here, but in the case study I choose to focus on the production of astronomical tables. Specifically, I consider the evolution of trigonometric tables, which are usually included in tabular work, and that are more directly responsive to evolution of computational techniques since they include no observations.

The key idea in this regard is that the introduction of cheap paper favored the "waste" of writing material, to the extent that the practitioner was no longer constrained to a limited area over which to compute and could include all the intermediate computations over the surface. This idea is quite intuitive once we consider the evolution of arithmetic techniques in the Islamicate world (or Latin Europe, for that matter) before and after the introduction of paper. Both the Islamicate word and Latin Europe originally developed arithmetic techniques designed to be deployed with tablets. The first was due to the fact that mathematical practice in Ancient Greece and Rome was often carried out on wax tablets (Bülow-Jacobsen 2012); the second because those same techniques arrived via Indian mathematicians, who did their calculations on dust tablets or boards. In the Islamicate world the association of arithmetical procedures with the dust board was so strong that what we call today Indian-Arabic numerals were referred to as "calculation on dust" (*hisāb-al-ghubār*), as does for instance Ṣāʿid al-Andalusī in his *Ṭabaqāt al-'Umam*.

The limitation of calculation with tablets is that these have a limited size, which the availability of cheap paper can circumvent. Although in the original case study I make a more indirect argument that assumed that the Iberian Peninsula followed the same developments in arithmetic practice as the Islamic East, I recently found a *single* manuscript that, in the course of going through different hands throughout the centuries, underscores this evolution—this is the *Ms. 189* in Christ Church College Library.[8] This is a 15th century copy of the Hebrew translation prepared by Moses ibn Tibbon in 1271 of Ceuti

[8] Since I cannot read Hebrew, I am relying on Pfeffer's (2017) thorough description of its contents.

mathematician Abu Bakr ibn Muhammad ibn Ayyash al-Hassar's 12th century *Kitāb al Bayān wa-l-tadhkār* (*The Book of Proof and Recall*), which contains elementary arithmetic topics with a pedagogical intent. Al-Hassar's original treatise deploys traditional erasure methods characteristic of dust arithmetic. I will not explain in detail how this system works—I refer the reader to the case study for that—but suffice to say that, added to the original copy, there is an undated Appendix titled "On the multiplication of integers by another method that is not from the book". This new method is the traditional pen-and-paper method of long multiplication.

The fact that the copier felt the need to add an appendix describing "another method" for long multiplication is very illustrative of the fact that pen-and-paper methods became handy when the computations became long and arduous. Before the introduction of paper, several methods existed to check previous work—Smith, for example, provides a review of the "method of nines" that several Arabic and Latin authors describe. (1958, 151 and ff.) But once the operations get large, even tremendously ingenious methods like the method of nines require a considerable amount of dedication and persistence. Once paper comes onto the scene, these methods become obsolete, and that is what we see in the historical record.

In the case study, I consider how these new arithmetic techniques changed the process of production of trigonometric tables. Some of these tables—such as those present in the Alfonsine Tables—had at least 360 entries, and included operations considerably more complex than long number multiplication. In this way, the introduction of paper favors further offloading of computational work, by allowing the computer to lay out some of the intermediate steps onto the writing surface. In doing so, we witness an expansion in intermediary documentation that changed how astronomy (or astronomical tabulation, at least) was made in the peninsula.

This concludes the within-case inference. Paper, I argued, induced in Medieval Iberia an expansion of intermediary documentation through the expansion of two kinds of documents: observation logs and computational offloading. In changing observational and computational practices, I think there is a case to be made that it had a considerable impact in Medieval Iberian astronomy as a whole. I argued that the main drivers of these new practices were some of paper's material properties, which motivated the actors to change their writing practices in acknowledgment of the new affordances provided by the material. I concurred with the received historiographical tradition that emphasizes the low price of paper. But I provided a series of intermediary mechanisms that made that hypothesis concrete. In the next and concluding subsection, I will consider how this within-case inference can tell us about the LNQA translation exercise.

7. Concluding remarks

In the second part of this chapter I have sketched how one might go about running an LNQA study of the impact of paper in Medieval Islamicate science. To do this, I: 1. Proposed an association in the form of an X Regularity (see R4), 2. Established a clearly defined scope of that regularity, 3. Produced a causal diagram that is meant to explain why that regularity obtains, and 4. I run within-case inference for a single case out of the five, illustrating how the causal model can be tested in the evidence provided by one of the cases.

Let me state some of the limitations of this work before I conclude. Earlier I argued that the Maghrebi case is something like a smoking-gun test for the causal hypothesis, in that it is the least likely case where the influence of paper should be seen. However, even if the Maghrebi case is shown to be conforming (which I hope I made a reasonable case it was), an LNQA study should continue the within-case inference to the other four cases within the scope. This is an important point because without some consideration of the other cases we would not be able to get an idea of the *strength* of the causal generalization, or even the strength of the association. In fact, further inquiry into nonconforming cases would provide opportunities to narrow the scope of the claim, thus better understanding the "special effects" that paper had in Medieval Islamicate science. Enriching the causal model would also give more insights into paper's special effects on Islamicate scientific practice. In future work, a more complete causal model should also include alternative explanations—for example, including the possibility that a rising of the standard of living allowed for much more liberal use of parchment.

Now onto the payoff. What have we achieved? Recall that my main objective was showing that case studies can indeed be good for producing general knowledge. I argued that at least part of the criticisms that dismiss that idea did not consider how methods in the social sciences use case studies to extract general knowledge. LNQA, I have tried to demonstrate, allows us to obtain generalized causal knowledge from the historical record, including the history of science.

In particular, we have been able to generalize a causal claim with an extraordinarily wide scope, spanning five centuries (9^{th}-14^{th}) across lands from the Atlantic to the Indian Oceans. This is a region and period with unprecedented significance for the history of science at large, where for the first time the Chinese, Persian, Greek, and Indian traditions coalesced into a system of knowledge and practice that would reinvigorate science in the centuries to come. Mathematics, optics, sociology or medicine would witness radically original achievements, and they would be disseminated and absorbed across the Afro-Eurasian continent for other civilizations and traditions to build upon. Understanding the

nature of this transformation is not only interesting in itself, but also a necessary part of a comprehensive history of science.

In this chapter, as well as in the chapters that follow, I have put the material infrastructure of production front and center in the factors that contribute to the historical development of scientific practice. The production of intermediary documents, far from being a marginal auxiliary to production, forms a basic part of the routine of the practice of science, and as such it greatly conditions what science can and cannot be done. To be sure, the development of papermaking will never be the crown achievement of the scientific endeavor like al-Khwārizmī's algebra or Ibn al-Haytham's optical theory. Like any infrastructure, paper sits in the pipes, the arteries, setting the space in which practice can be exercised, and on top of practice, theory produced. It is for this reason that we should expect changes in the infrastructure of production to be also of great philosophical relevance, by revealing part of the causal structure that constrains and affords the changes that we see in the historical record.

In the next chapter, I will take up the "smoking gun" case—medieval Iberia—in earnest, and ask how the introduction of paper could have contributed to the expansion of intermediary documentation, with a focus on the 11th to 13th centuries. By providing some reasons to believe that the generalization holds in that least likely case, I hope to make the case for the broader generalization across the medieval Islamicate world. After that, I will consider the other horn of the generalization, and show how the case study of medieval Iberia shows a way to avoid cherry-picking in case study work in the history of sciece.

References

Afsaruddin, A. "Caliphate." Encyclopedia Britannica, December 5, 2025. https://www.britannica.com/place/Caliphate.

Bennett, Andrew. 2004. “Process Tracing and Causal Inference.” In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by E. Henry and David Collier. Rowman & Littlefield.

Berggren, J. L. 2016. *Episodes in the Mathematics of Medieval Islam*. Springer. https://doi.org/10.1007/978-1-4939-3780-6.

Birdsall, Nancy M., Jose Edgardo L. Campos, Chang-Shik Kim, et al. 1993. *The East Asian Miracle: Economic Growth and Public Policy (Vol. 1 of 2): Main Report (English). A World Bank Policy Research Report*. World Bank Group. http://documents.worldbank.org/curated/en/975081468244550798.

Blair, Ann M. 2010. *Too Much to Know: Managing Scholarly Information before the Modern Age*. Yale University Press.

Bloom, Jonathan. 2001. *Paper before Print: The History and Impact of Paper in the Islamic World*. Yale University Press.

Bloom, Jonathan M. 2005. “Silk Road or Paper Road?” *The Silk Road* 3 (2): 21–26.

Boucoyannis, Deborah. 2021. *Kings as Judges: Power, Justice, and the Origins of Parliaments*. Cambridge University Press.

Braumoeller, Bear F., and Gary Goertz. 2000. “The Methodology of Necessary Conditions.” *American Journal of Political Science*, 844–58.

Bülow-Jacobsen, Adam. 2012. “Writing Materials in the Ancient World.” In *The Oxford Handbook of Papyrology*, 1st ed., edited by Roger S. Bagnall. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199843695.013.0001.

Carruthers, Mary J. 1992. *The Book of Memory: A Study of Memory in Medieval Culture*. no. 10. Cambridge University Press.

Cartwright, Nancy, and Jeremy Hardie. 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford University Press.

Cartwright, Nancy, Brendan Kelters, Eileen Munro, Likhwa Ncube, and John Pemberton. 2006. “Can I Make This Policy Work Here?” Unpublished manuscript.

Coleman, James S. 1986. “Social Theory, Social Research, and a Theory of Action.” *American Journal of Sociology* 91 (6): 1309–35.

Constable, Olivia Remie. 1994. *Trade and Traders in Muslim Spain: The Commercial Realignment of the Iberian Peninsula, 900-1500*. Cambridge Studies in Medieval Life and Thought, 4th ser., 24. Cambridge University Press.

Déroche, François, Annie Berthier, Deke Dusinberre, David Poole Radzinowicz, and Muhammad I. Waley. 2015. "Islamic Codicology: An Introduction to the Study of Manuscripts in Arabic Script." *(No Title)* (London).

Deyo, Frederic C. 2020. *Beneath the Miracle*. Reprint 2020. UC Press Voices Revived. University of California Press.

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *COMPARATIVE POLITICS-NEW YORK-* 30: 127–46.

Eisenstein, Elizabeth L. 1980. *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early Modern Europe, Volumes I and II*. Cambridge University Press. https://doi.org/10.1017/CBO9781107049963.

Findley, Michael G., Kyosuke Kikuta, and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* 24 (1): 365–93.

García Díaz, I., and J. A. Montalbán Jiménez. 2005. "El Uso Del Papel En Castilla Durante La Baja Edad Media." *Actas Del VI Congreso Nacional de Historia Del Papel En España: Buñol* (Buñol).

Geddes, Barbara. 2010. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Nachdr. Analytical Perspectives on Politics. Univ. of Michigan Press.

Goertz, Garry, and Stephan Haggard. 2025. "Large-N Qualitative Analysis: Causal Generalizations and Multimethod Research." Unpublished manuscript.

Goertz, Gary. 2020. *Social Science Concepts and Measurement: New and Completely Revised Edition*. Princeton University Press.

Goitein/Lassner. 2001. *Mediterranean Society, An Abridgment in One Volume*. Edited by S. D. Goitein and Jacob Lassner. University of California Press. https://doi.org/10.1525/9780520354227.

Gutas, Dimitri. 2012. *Greek Thought, Arabic Culture: The Graeco-Arabic Translation Movement in Baghdad and Early'Abbasaid Society (2nd-4th/5th-10th c.)*. Routledge.

Ḥamawī, Yāqūt ibn-ʿAbdullāh al-Rūmī al-. 1861. *Dictionnaire Géographique, Historique et Littéraire de La Perse et Des Contrées Adjacentes: Extrait Du Mo'djem El-Bouldan de Yaqout, et Complété à l'aide de Documents Arabes et Persans Pour La Plupart Inédits*.

Hirschler, Konrad. 2016. “From Archive to Archival Practices: Rethinking the Preservation of Mamluk Administrative Documents.” *Journal of American Oriental Society* 136 (1): 1–28.

Hodgson, Marshall G. S. 1974. *The Venture of Islam. 1: The Classical Age of Islam*. Paperback ed., 17. pr. Univ. of Chicago Press.

Hoffman, Adina, and Peter Cole. 2011. *Sacred Trash: The Lost and Found World of the Cairo Geniza*. 1st ed. Jewish Encounters. Nextbook : Schocken.

Homola, Jonathan, Miguel M. Pereira, and Margit Tavits. 2020. “Legacies of the Third Reich: Concentration Camps and out-Group Intolerance.” *American Political Science Review* 114 (2): 573–90.

Ibn Khaldūn. 2015. *The Muqaddimah: An Introduction to History*. Edited by N. J. Dawood. Translated by Franz Rosenthal. With Bruce B. Lawrence. Princeton University Press.

Karabacek, Joseph von, and Don Baker. 2001. *Arab Paper*. Archetype Publications.

Kelanic, Rosemary Ann. 2012. *Black Gold and Blackmail: The Politics of International Oil Coercion*. The University of Chicago.

King, David A. 2018. “Astronomy in Medieval Jerusalem.” Unpublished manuscript.

Kremer, Alfred Freiherr von. 1920. *The Orient Under the Caliphs*. University of Calcutta.

Kurtz, Marcus J. 2000. “Understanding Peasant Revolution: From Concept to Theory and Case.” *Theory and Society* 29 (1): 93–124.

Levi-Provençal, Évariste. 1934. “Un Manucrit de La Bibliothèque Du Calife Al-Hakam II.” *Hespéris* (Paris), 198–200.

Levitsky, Steven, and Lucan Way. 2022. *Revolution and Dictatorship: The Violent Origins of Durable Authoritarianism*.

Levy, Jack S., and William R. Thompson. 2011. *Causes of War*. John Wiley & Sons.

Mahoney, James. 2012. “The Logic of Process Tracing Tests in the Social Sciences.” *Sociological Methods & Research* 41 (4): 570–97.

Mazur, Kevin. 2021. *Revolution in Syria: Identity, Networks, and Repression*. Cambridge University Press.

McCluskey, Stephen C. 2000. *Astronomies and Cultures in Early Medieval Europe*. Reprinted. Cambridge Univ. Press.

Millás Vallicrosa, J. M. 1943. *Estudios Sobre Azarquiel*. Consejo Superior de Investigaciones Científicas–Instituto Miguel Asín–Escuelas de Estudios Árabes de Madrid y Granada.

Miller, Michael K. 2021. *Shock to the System: Coups, Elections, and War on the Road to Democratization*.

Morelon, Régis. 2019. "Eastern Arabic Astronomy between the Eighth and the Eleventh Centuries." In *Encyclopedia of the History of Arabic Science*. Routledge.

Paul, Jürgen. 2015. "Archival Practices in the Muslim World Prior to 1500." In *Manuscripts and Archives: Comparative Views on Record-Keeping*, vol. 11. De Gruyter.

Pellat, Charles, and Reinhart Pieter Anne Dozy. 1961. *Le Calendrier de Cordoue*. Vol. 1. Brill Archive.

Pfeffer, Jeremy. 2017. "Moses Ibn Tibbon's Hebrew Translation of al-Hassar's Kitab al Bayan." *Convergence*, ahead of print. https://doi.org/10.4169/convergence20170501.

Pietsch, Wolfgang. 2016. "Two Modes of Reasoning with Case Studies." In *The Philosophy of Historical Case Studies*. Springer.

Pitt, Joseph C. 2001. "The Dilemma of Case Studies: Toward a Heraclitian Philosophy of Science." *Perspectives on Science* 9 (4): 373–82. https://doi.org/10.1162/106361401760375785.

Saliba, George. 1987. "Theory and Observation in Islamic Astronomy: The Work of IBN AL-SHĀTIR of Damascus." *Journal for the History of Astronomy* 18 (1): 35–43. https://doi.org/10.1177/002182868701800102.

Saliba, George. 1995. *A History of Arabic Astronomy: Planetary Theories during the Golden Age of Islam*. Vol. 2. NYU Press.

Saliba, George. 2007. *Islamic Science and the Making of the European Renaissance*. The MIT Press. https://direct.mit.edu/books/monograph/2074/Islamic-Science-and-the-Making-of-the-European.

Samsó, Julio. 2020. *On Both Sides of the Strait of Gibraltar: Studies in the History of Medieval Astronomy in the Iberian Peninsula and the Maghrib*. Handbook of Oriental Studies, volume 144. Brill.

Sauer, Tilman, and Raphael Scholl, eds. 2016. *The Philosophy of Historical Case Studies*. Vol. 319. Boston Studies in the Philosophy and History of Science. Springer International Publishing. https://doi.org/10.1007/978-3-319-30229-4.

Schoeler, Gregor. 2009. *The Genesis of Literature in Islam: From the Aural to the Read*. The New Edinburgh Islamic Surveys. Edinburgh University press.

Shatzmiller, Maya. 2018. “The Adoption of Paper in the Middle East, 700-1300 Ad.” *Journal of the Economic and Social History of the Orient* 61 (3): 461–92. https://doi.org/10.1163/15685209-12341456.

Smith, David Eugene. 1958. *History of Mathematics*. Vol. 2. Dover.

Staniland, Paul. 2021. “Leftist Insurgency in Democracies.” *Comparative Political Studies* 54 (3–4): 518–52.

Thaʾālibī, ʿabd al-Malik ibn Muhammed al-. 1968. *The Book of Curious and Entertaining Informaion (Latāʾif al-Maʾārif)*. Translated by C. E. Bonsworth. Edinburgh University Press.

Valls i Subirà, Orioll. 1978. *La Historia Del Papel En España, Siglo X-XIV*. Vol. 1. Empresa Nacional de Celulosas.

Vallvé Bermejo, Joaquín. 1980. “La Industria En Al-Andalus.” *Al-Qantara* 1 (1–2): 209–42.

Van Evera, Stephen. 1997. “What Are Case Studies? How Should They Be Performed.” *Guide to Methods for Students of Political Science*, 49–88.