Replication and the search for the laws in the geographic sciences

Peter Kedron[1] and Joseph Holler[2]

[1] Assistant Professor, School of Geographical Sciences and Urban Planning, Arizona State University
[2] Assistant Professor, Department of Geography, Middlebury College

Forthcoming as Kedron, P., Holler, J. Replicability in geographic science and the search for laws of geography. *Annals of GIS.*

Abstract: Replication is a means of assessing the credibility and generalizability of scientific results, whereby subsequent studies independently corroborate the findings of initial research. In the study of geographic phenomena, a distinct form of replicability is particularly important - whether a result obtained in one geographic context applies in another geographic context. However, the laws of geography suggest that it may be challenging to use replication to assess the credibility of findings across space and to identify new laws. Many geographic phenomena are spatially heterogeneous, which implies they exhibit uncontrolled variance across the surface of the earth and lack a characteristic mean. When a phenomenon is spatially heterogeneous, it may be difficult or impossible to establish baselines or rules for study-to-study comparisons. At the same time, geographic observations are typically spatially dependent, which makes it difficult to isolate the effects of interest for cross-study comparison. In this paper, we discuss how laws describing the spatial variation of phenomena may influence the use of replication in geographic research. Developing a set of shared principles for replication assessment based on fundamental laws of geography is a prerequisite for adapting replication standards to meet the needs of disciplinary subfields while maintaining a shared analytical foundation for convergent spatial research.

Keywords: Replication, Spatial Heterogeneity, Spatial Autocorrelation, Laws of Geography

Introduction

The pursuit of new scientific knowledge operates under the assumptions that nature follows consistent rules within equivalent contexts and that our knowledge grows as scientists test those rules under contextual variations to determine if and how they persist or change. Rules that are regularly supported by evidence from many different studies conducted under different conditions may be elevated to the status of laws – general, synthetic descriptions inferred from systematic observations that hold under specified conditions.

Whether geography is a law-seeking discipline and what role laws play in geographic research have been topics of recurrent debate throughout the post WWII period. Hartshorne (1939, 1954, 1955) and Schaefer (1953) famously debated the subject, preceded by Ackerman's (1945) call to systematize the discipline, and immediately succeeded by Bunge’s (1966) and Harvey's (1969) attempts to strengthen the scientific foundations of geographic explanation. With these early works as recurrent touchstones, geographers continued to debate the position of laws in the discipline along two lines of argument. First, geographers exchanged ideas about the characteristics a law would need in order to be classified as a geographic law (Bunge, 1966; Hay, 1979, 1985; Sack 1972, 1973, 1980). Second, geographers debated whether the uniqueness of places precludes the possibility of discovering laws of geography altogether (Bunge, 1966; Guelke, 1977; Lewis, 1972). This second line of argument later fused with the critiques of radical, humanist, and feminist geographers (see Kitchin, 2006) who interrogated the objectivity, ontology, and epistemology behind law-seeking, positivist geography. However, geography’s disciplinary focus on the uniqueness of place and the objectivity or subjectivity of laws shifted attention away from another fundamental question—if geographic laws do exist, how are they discovered?

Across the sciences, replication is key to the identification of laws. In a replication study, a researcher repeats the procedures of an existing study while intentionally changing one or more research parameters (e.g. the study site, study population, confounds, etc.). Replication studies may be motivated by law seeking or a variety of other factors, including the practical application of existing studies to new contexts or the integration of established studies into convergence research. Regardless of motivation, each replication study may provide opportunities to test the validity and generalizability of the original study's claims.

In this paper, we examine the connection between replication, the existing laws of geography, and the pursuit of new scientific knowledge in our discipline. To do so, we examine how the characteristics of geographic phenomenon complicate the accumulation of the empirical support needed to establish laws within the discipline. We frame our discussion around the two most widely cited laws of geography. First, Tobler's law (Tobler, 1970) famously states that, "everything is related to everything else, but near things are more related than distant things". Second, Anselin's (1989) proposed law of spatial heterogeneity highlights that phenomena vary in space. We focus our examination on these laws because they are fundamental to understanding the complexity and variability of geographic systems, and are characteristics that complicate the establishment of the thresholds of evidence needed to establish belief and the use of replication to establish the credibility of scientific theories and laws in geography.

We have organized the reminder of this paper as follows. In the following section, we briefly present the historic treatment of laws in geography and highlight the limited explicit treatment of replication. We next define scientific laws, outline criteria for their identification, and discuss the role replications play in establishing the credibility and range of laws. We then present some of the ways spatial heterogeneity and spatial dependence complicate the use of replication in geography, before concluding with a discussion of how ongoing research might further our initial analysis.

Replication in Geographic Sciences

Geographers at the forefront of the quantitative revolution (e.g. Bunge, 1966; Chorley & Haggett, 1965; Harvey, 1969) developed a preliminary framework for discovery of geographic laws by specifying robust standards of empirical support and formal conventions for statements of law. Their work built on and tailored contributions by philosophers of science (Bergson, 1950; Bergmann, 1957; Popper, 1959; Braithwaite, 1960; Hempel, 1965) to the geographical sciences. With regard to standards of empirical support, a clear emphasis on the need for repeated demonstration of a proposed relationship in different contexts emerged – *a need for replication*. Focusing on laws as descriptions of spatial pattern, Bunge (1966) highlighted that a single test of a pattern relationship is not sufficient to turn that relationship into a law, and that replication will be needed to establish the generality of the relationship. Golledge and Amedeo (1968) reinforced this position highlighting the need for repeated observations in varied circumstances and locations before a law can emerge. Harvey (1969) took a similar position at several points during his discussion of scientific explanation and laws and theories in *Explanation in Geography*. The need for replication was acknowledged in nearly every early text on law-seeking in geography, confirming its fundamental importance. And yet, geographers have not developed the mechanisms through which replication should—or cannot— function to discover and develop geographic law.

Neither the advocates nor the critics of law-seeking in geography gave detailed attention to replication even though both groups shared an understanding of replication as an essential part of the scientific method. Advocates of law-seeking devoted little effort to the development of formal frameworks and methodologies for replication studies; and critics of law-seeking only occasionally took direct aim at the lack of replicability in geography. Early criticism of law-seeking geography by Guelke (1971, 1978) and Gregory (1978) could have brought replication to the forefront as both authors point out that geographers seeking laws often failed to rectify their claims with contradictions found in repeated, empirical observations. Sayer (1992) made the fundamental point that observing regular associations between events is not enough to justify a law. Sayer drew a clear distinction between instrumental laws that refer to regularities among events but not causal mechanisms, and causal laws that identify those mechanisms. The ways in which these two types of laws operate in open and closed systems remains undeveloped as related to replication in geography.

Independent of the debates on law-seeking geography, the role of replication in geographic research needs deeper study simply because determining the degree to which a replication study supports prior results is often not apparent. Indeed, the emphasis radical and human geographers place on the need to contextualize and interpret place-based variation in phenomena (see Peet, 1999) in their critiques of law-seeking suggests the need for a careful investigation of replication in space and time. For example, when are the results of two studies conducted in different location similar enough to be considered supportive of one another? How should context be accounted for in such a comparison? How should the geographic distribution of replication studies be designed and later weighted when assessing the credibility of any claimed effect?

In this paper, we emphasize the dimensions of geographic system complexity related to the first two laws of geography as the first step to understanding the structure and complexity of geographic systems with regard to the design and interpretation of replication studies. As such, we place our emphasis on space and complexity of spatial patterns. However, we recognize that other related dimensions of complexity also exist in geographic systems and will confound the design and interpretation of replication studies. These dimensions include temporal autocorrelation, temporal non-stationarity, and processes or feedbacks across multiple scales in nested or interconnected systems. Each of these dimensions of complexity in geographic systems will require additional work to investigate and formalize their implications for replicability and replication studies, while the following discussion remains focused on spatial autocorrelation and heterogeneity.

Davies (1968) provided the most useful explication of replication from the positivist geographic perspective. In a paper examining the predictions of central place theory, Davies simultaneously: distinguished between research data, techniques, and results; acknowledged the complications spatial-temporal variations create for replication attempts; and presented two empirical examples of early attempts at replication. Each of these elements are fundamental to present approaches to replication across the sciences[1]. Notwithstanding the importance of Davies' paper, surprisingly few geographers have cited the work or followed his example in the intervening 50 years[2]. Formal published replications of geographic research remain scarce, and detailed assessments of how replications may or may not

[1] See Schmidt (2009) for a highly cited presentation of the importance of distinguishing between the different parts of a study for replication, chapter 5 of NASEM (2019) for a discussion of the fundamental role of variation in assessing replication, and Roesch and Rougier (2020) as an example of the establishment of *ReScienceX* an open journal of peer-reviewed replications.

[2] See Waters (2020) for a discussion of a few exceptions.

function in geographic research remain even scarcer (see Konkol et al., 2019; Nust et al., 2020; Stigell & Schantz, 2011; Wainwright, 2020).

The past decade has seen replication and reproduction move forward on research agendas across the sciences (NASEM, 2019), and in the past five years these topics have garnered increasing interest within the geographical sciences (Brunsdon, 2016; Kedron et al., 2019; Brunsdon & Comer, 2020; Goodchild et al., 2020; Kedron et al. 2020). However, much of this attention has focused primarily on reproduction rather than replication, with a particular emphasis in the literature on computational forms of reproducibility (see Rey, 2009; Konkol et al., 2019; Nust & Pebesma, 2021). Reproductions repeat an original analysis using the same data and procedures in order to assess the original result and internal validity of the study. As a practical matter, independent researchers can most easily undertake replication studies when the researchers who conducted the original study provide sufficient information about their procedures. Including a full record of the provenance of a result allows others to understand what was done and how (Tullis and Kar, 2021). A full provenance record enables independent researchers to plan replication studies by intentionally altering research parameters and decreases uncertainty when comparing replication results with the original study. Ideally, researchers accompany their record of research procedures with a research compendium encompassing the code, data, processing environment, and metadata for the study (Wilson et al 2021, Nust & Pebesma, 2021). While generating and sharing this record may appear trivial, a growing body of literature catalogs the many complications (e.g., data privacy, multi-user environments, equipment/operator uncertainty) that researchers may encounter when they use varied and complex computational methods in their research (see Allison et al., 2018; Millman et al., 2014; NASEM, 2019). Substantial investments are being made into infrastructure to improve scientific reporting practices and facilitate data and code sharing to make research more computationally reproducible and replications easier and more meaningful to pursue (Wang, 2010; 2016, Richardson, 2019)

Laws and Replication in the Pursuit of Scientific Knowledge

In this paper, our focus is on replication and with a single primary motivation—the desire and need to preserve a core evaluative mechanism of science so we can continue to produce credible descriptions and explanations of phenomena. Reproducibility is an essential prerequisite for useful replication studies because it enables researchers to purposefully isolate and alter selected research parameters. If the original study is not fully reproducible and a replication study produces contradictory results, there will be uncertainty in attributing the cause to either a misspecification of procedures in original research, or to errors or contextual limitations with the theorized effects. While reproducibility has many benefits—transparency, public trust, and facilitating convergence research, among others—our primary interest here is in reproducibility and reproduction as a first step toward conducting replication studies with meaningful results for testing the generalizability of scientific knowledge. To date, published studies in geography have given little explicit focus to this motivation, further reinforcing Kitchin's (2006) critique that positivistic geography has yet to deeply reflect on its philosophical underpinning. Throughout the remainder of this paper, we begin to address this gap in the literature by explicitly linking the pursuit of scientific knowledge within geography to replication and framing that endeavor within the widely cited first- and second-laws of geography.

The characteristics of laws

A scientific law is a synthetic statement that describes how some phenomenon will behave under a set of conditions. Laws describe regular associations, modes of behavior, or patterns that are relatively stable and apply to all the phenomena they describe (Castree, 2005). Laws have three key features that distinguish them from other forms of synthetic statements and set criteria for their identification and assessment. Laws must be (1) general statements about factual truths, (2) empirically supported, and (3) integrated into theory (Braithwaite, 1960; Hempel, 1965; Golledge & Amedeo, 1968; Harvey, 1969).

First, laws are general statements about all instances of a kind rather than statements about an individual instance. For example, the statement that 'elevation exhibits uncontrolled variance in Vermont' is not a law because it applies only to the specific instance of elevation in a single state. In contrast, the proposed second law of geography that *'geographic variables exhibit uncontrolled variance*' is a statement applicable to any phenomena that can be represented as a geographic variable. The statement similarly makes a universal claim that is not bounded by location or time.

Second, as synthetic statements, the validity of laws cannot be established by analyzing the definitions of their concepts, but rather have to be discovered and verified empirically through experience, observation, or experimentation. In this sense, laws are summaries of stable relationships that have been repeatedly observed. However, because laws make statements about all instances of a kind, the empirical evidence in support of a law is always incomplete. Continuing the example above, observing variation in elevation in Vermont may be sufficient to establish the truth of the specific statement of fact that elevation exhibits uncontrolled variance in that state. In contrast, the validity of the proposed second law of geography depends on empirically observing uncontrolled variance in all instances of geographic phenomena. As we are unable to observe all possible instances of a phenomena across space and time, no accumulation of supporting empirical evidence can conclusively verify the law. However, a single instance of a geographic phenomenon not exhibiting uncontrolled variance may bring the validity of a second law into question. As Popper (1959) argues, laws are conclusively falsifiable, but not conclusively verifiable.

Recognizing that strict adherence to qualities of universality would lead to the elimination of laws across the sciences, both geographers (see Harvey 1969) and philosophers of science (see Nagel 1961) generally relax the requirement that a scientific law hold for all instances under all conditions[3]. Instead,

[3] Even Popper made a distinction between the logic of falsifiability and its application, recognizing that no set of observations are entirely free from error and potential unobserved confounds.

geographers make a practical decision to treat some statements as if they were universally true knowing that those statements will never be (can never be) shown to be universally true - a position Harvey terms 'methodological universality'. This compromise creates the additional need for demarcation criteria that identify when enough evidence has been gathered across a sufficient range of conditions to justify this position. There is no general agreement on what these criteria should be. However, as Harvey (1969) points out, lacking these criteria and making the underlying assumption of universality does not diminish the utility of a law in situations in which it has been supported, and only becomes problematic when we use the law to make inferences beyond the range of conditions for which evidence has supported it. Disagreement about demarcation criteria may stem, at least in part, from the different approaches adopted in different sub-fields of the geographical sciences and the nature of the phenomena those fields typically investigate.

Irrespective of demarcation criteria, universality may also be relaxed in different ways. Most directly, the author of a law can simply limit the domain over which a law holds through the specification of additional conditions. Another way to relax the universality criterion of a law is by specifying the law probabilistically (Jones, 1956)–to describe non-deterministic relationships that have only have a certain chance of occurring for a certain class of phenomenon. Probabilistic laws are assessed by collecting empirical evidence across a large number of instances of a class and the number of occurrences is then evaluated against the number predicted by the law. In the geographic case, if a statistical law is tested in many locations, then the frequency across the set of locations would be used to assess the law. As long as the probability proposed by the law was less than one, the law would be supported even though some locations did not demonstrate the proposed outcome.

Third, to be considered a law, a statement should be part of a theoretical system and supported by other components of that system (Harvey, 1969). The relationship between a law and a theory is an important one because laws are themselves only descriptions of phenomena in the world. Laws do not explain how or why phenomena operate as they do. For example, the second law of geography does not make any statement as to why geographic variables like elevation exhibit uncontrolled variance. Explanations of phenomena are given by theories, and it is through the explanatory structure of theories that we are able to create testable, falsifiable hypotheses. Repeatedly testing theoretically informed hypotheses ultimately contributes empirical support for a generalizable law. For example, linking the second law of geography to theories of geomorphology could produce localized estimates of elevation that can be tested through observation.

Assessing the credibility of laws through replication

Replication studies answer the question of whether a result can be found again in a broader set of study contexts, thereby directly addressing two of the criteria for identifying a law - the level of empirical support for the law and the universality of its description.

A valid law specifies a predictable relationship that researchers can develop into a falsifiable hypothesis for empirical testing and evaluation. Belief in the law is founded in empirical studies that confirm hypotheses derived from the law. Such empirical studies should be replicable. The scientific community can adjust belief in the law and its theorized relationship based on the outcomes of new replication studies, our initial confidence in the hypothesized relationship, and the extent to which the replication appropriately tested that relationship. Repeated across a series of studies, the information gleaned from each replication progressively transforms our belief (Earp & Trafimow, 2015), and this process could be formalized with Bayesian statistics (Nichols et al., 2021). Even in the presence of moderate researcher bias, consistent supportive evidence from high-quality replications can quickly increase confidence in the feasibility and veracity of a hypothesis (Coffman & Niederle, 2015). Conversely, replications that produce contradictory evidence can diminish our degree of belief or lead us to expand the set of conditions we use to limit the law.

Replication research is often complicated by a number of factors. Ideally, a researcher testing a proposed law with a replication study will be able to specify the complete set of conditions under which the relationship defined by the law is expected to hold, control for factors that could potentially confound the relationship, and reliably measure the variables internal to the law. Complex systems with uncontrollable or unknown confounds are typically less amenable to replication (NASEM, 2019). However, complexity and a lack of control are key characteristics of the systems studied by geographers. The analogous case of laws within the social sciences is instructive. Kincaid (1990) argues that laws in social systems may be limited to idealized forms that operate only under *ceteris paribus* conditions, but also that such laws can nonetheless be assessed. He offers six testing practices for assessing laws when the specification of confounding factors is incomplete, of which we will highlight the three most relevant to replication studies. These practices hinge on the idea of testing relations within and around the limits of the conditions tied to a proposed law. First, studies may repeat testing within the narrow range of existing cases in which the law's conditions are satisfied and can be confirmed. This uses replication to confirm that the theorized results were not spurious within the specified context. Second, studies may demonstrate that different deviations from the required conditions have little impact on the law. This uses replication to build our confidence in the generalizability of the law to progressively broader contexts. Third, inductive reasoning may be applied to a collection of studies to assess whether the predicted relationship of a law is more accurate as complex social conditions more closely match the law's theorized conditions. This uses replication to verify how dependent the supposed law is on a set of contextual parameters. If important contextual parameters were never specified, then contradictory replication results may help to discover and specify additional *ceteris paribus* conditions.

The usefulness of a replication as a test of a law depends on a researcher's ability to compare the result of that study with prior results or to the relationship defined by a law or theory. While there is no single approach to determining the consistency of study results, there is agreement that any approach needs to consider the proximity of the results, the degree of uncertainty associated with their measurement, and the variability of the system being studied (NASEM, 2019). Defining the proximity of the result of a replication to the results of other studies and the relationship proposed by a law will be shaped by how a law is formulated. Laws can be presented in a number of ways and identifying whether a law makes a statement about the direction, magnitude, and the functional form of an association are all key to assessing the proximity of results. For example, Tobler's first law defines a directional relationship that is a function of distance, but does not provide enough information to assign an expected magnitude or specific functional form[4]. In contrast, physical laws governing geomorphological processes or those connected to central place theory present precise functional relationships that prescribe calculable magnitudes under given conditions. If a law only suggests a directional relationship, evidence of consistency in direction across studies may raise the credibility of the law. In contrast, simply finding matching directional relationships between two studies would not be enough to support a law that prescribes a specific magnitude of effect.

Even when a law is precisely stated, there will always be some level of uncertainty in our measurements of the key phenomena. For example, if a replication study finds a magnitude of effect that differs from an original study, but both effects fall within each other's confidence intervals, we may determine that the second study has replicated the first. Our level of uncertainty will not only vary from study to study, but also from system to system. When assessing replications, we must also account for the variability of systems. However, as geographers we recognize that the variability within and between systems is also intrinsically connected to the laws of spatial dependence and heterogeneity, confounding our interpretation of both original empirical studies and their replications.

The First and Second Laws of Geography – Confounds of Replication and the Search for Laws in the Geographical Sciences

The first and second laws of geography create two paradoxes concerning the use of replication as a means of collecting the empirical evidence needed in the search for geographic laws. A first paradox concerns the tension between Tobler’s law and Anselin’s principle when describing geographic phenomena. Tobler's law predicts similarity among events proximate in space that leads to some expectation of similar or confirming study results among geographically proximate replications. Conversely, Anselin's principle tells us to expect geographic phenomena to vary across space giving us reason to doubt that the relationship proposed by a law should hold in all locations. A lack of clarity about where to expect a law to hold or to vary in consistency is a fundamental challenge to the conduct and assessment of replication studies in the discipline. For example, Christaller’s central place theory may

[4] Reformulating the law as Goodchild (2004) does as, "for every geographic variable a function of location z=f(x) there exists some distance *d* below which covariance is monotonically increasing" makes the absence of a specified magnitude clear. This lack of specificity is also likely one of the reasons the law is so widely applied.

predict the spatial relationship of patterns of human settlement locations and sizes within the Champlain Valley of western Vermont, but the relationship quickly breaks down as one moves eastward through the Green Mountains, where the spatial heterogeneity of terrain confounds central place theory's assumption of an isotropic plane.

The heterogeneity principle similarly implies that relationships proposed by a geographic law will depend on the geographic scale and extent of an individual study. Generalizations made at one scale may not hold at another scale (Haggett et al. 1965), implying that geographic laws must specify the geographic contexts (including both the extent and support) in which the predicted relationship is expected to hold. If the spatial context of a law is not specified, a replication study's failure to confirm the relationship predicted by a law may be explained by differences in the scale of the study or simply the aggregation of empirical observations into geographic units of analysis. When replications across scales fail to support a law, belief in the relationship may stand if the law is refined to be a statement operating at a specific scale or geographic context. Continuing with the example of central place theory, statements of law derived from that theory should include the extent of region(s) studied and the scale of empirical observation and quantitative abstraction of human settlements.

Applying Tobler's law to the question of expectations, we may believe that replication studies conducted in geographic proximity to an original study are more likely to support the relationship proposed by a law than those conducted farther away. For example, replication studies of central place theory are more likely to confirm the predicted spatial relationship when studying regions approximating isotropic planes proximate to southern Germany. Indeed, exceptions to the theory begin to accumulate in attempts to apply the theory to distant former colonies, most of which are still developing countries. In such distant places, the economic interactions between people and settlements and between settlements and environmental resources are sufficiently different to produce anomalies, e.g. primate gateway cities (Rose 1966, Sjøholt 1984). Recognizing where a study is conducted and if a change in geographic context merits identifying a study as a replication is therefore important to judging how strongly a result supports a law and how much we should change our beliefs. Carefully executed replications across distant and diverse geographic contexts can create a body of evidence that, even if difficult to directly compare, may lead to the identification of key conditions of potential laws. Conversely, failing to account for the location of replications could lead to premature belief in the generalizability of a law. The sensitivity of replications to geographic context again suggests a clear need for a careful enumeration of the conditions mediating the relationship proposed in any prospective law of geography, as well as the criteria for demarcation of thresholds for establishing universality.

A second paradox concerns ambiguity in the specifications of the first and second laws. Any prospective geographic law should likely contain statements about the probable effects the first and second laws will have on the relationship being proposed, but the first and second laws themselves contain few details about when or how they will impact other relationships. With such general specifications, geographers of all epistemological dispositions routinely find evidence confirming that near things are indeed more closely related than distant ones and that phenomena vary in space. However, the degree to which phenomena vary in space or the strength of their relationship with near things varies with the phenomena under investigation. To account for these factors in the design of replication studies, it would be important to identify which aspects of a proposed law are expected to vary across space and account for those aspects in the specification of thresholds for determining universality. That task is complicated by the fact that laws can be formulated in different ways.

For example, specifying either the magnitude or probability of the relationship defined by a law can change how replications support or refute the law. If a law specifies both the direction and magnitude of a relationship a successful replication may partially support the law if the direction of the relationship holds, but the first and second laws amplify or dampen the magnitude of the relationship. Tobler’s law is a complicated confound in this case because it creates two effects. First, Tobler's law suggests the accuracy of our estimates of a law's proposed effect will likely be impacted by the strength of the relationships among observations in space. However, the strength of spatial relationships and the resulting confound are not often known, which makes it difficult to control for their effects during sample design, estimation, and inference. One way to account for the impact of spatial autocorrelation on the estimate of an effect is to widen the uncertainty estimate that accompanies that effect estimate. However, this may have the counterintuitive consequence of making it easier to find evidence of replication if comparisons between studies are based on the overlap of widened confidence intervals. Second, Tobler's law suggests that it will always be difficult to isolate the relationship proposed by a law from surrounding confounds that are often unknown. This effect makes it difficult to identify and include additional conditions on a law.

These issues are amplified if a law is formulated probabilistically. In this case, we would not necessarily know if failure to observe the relationships in any particular instance was attributable to the law being false, or simply the result of its probabilistic nature. Furthermore, if the second law of geography applies to probability of the relationship itself, we may not expect the prior probability of the law holding to be the same in all locations. While the heterogeneity principle states that phenomena vary in space, it does not specify how a phenomena or the probability of their occurrence will be distributed across spaces. However, the distribution of a phenomena has important implications for the use of replication and the search for laws. For example, if the relationship defined by a law has a magnitude that is normally distributed across space, many common statistical tests will likely provide reliable estimates of result proximity. However, if relationship magnitude instead follows a power law distribution, the relationship would exhibit uncontrolled variance across space and lack a characteristic mean. If this is the case, it would be difficult to make pair-wise comparisons of relationships using replications without knowing where the samples lie within the distribution. Any assessment of a relationship proposed by a law would need to rely on a larger systematic assessment of many replications conducted at many locations.

In this section, we have illustrated that the geographic laws of spatial dependence and heterogeneity raise significant confounding problems for development of new laws of geography through replication studies due to their influence on the complexity and variance within and between geographic systems. They must be accounted for in the specification of a law's geographic context, in the measurement of direction and magnitude of the predicted relationship, in specifying the threshold for determining and believing universality, and in designing replication studies. Given a proposed relationship based on an empirical study, replication studies are required to determine the degree of influence of the first two laws on—and the universality of—the proposed relationship. Furthermore, the replication studies must be able to hold the research parameters constant while varying only the geographic scale or location. It is not possible nor is it necessary to describe the full complexity and suite of confounding variables of the systems of study in social science (Kincaid 1990). However, geographers may use more precise replication studies across space, time, and scale to more thoroughly control for the influence of the first two laws of geography in pursuit of discovering novel spatial relationships and more precisely specifying the context of their universality.

Discussion

We have made a broad case for the role replication can play in the search for laws and scientific knowledge in geography and for the related need to explore how two established laws of geography may confound that process. In our discussion, we expand upon the confounding laws of geography and briefly present other typologies for classifying laws. We remark on why so few geographic laws have been discovered, and discuss how challenges of replication and the discovery of generalizable scientific knowledge in the geographical sciences.

We should consider the connection between replication and law in light of Guelke's (1978) critique that geography has developed few testable laws of general application. While the laws at the center of this paper undergird nearly all forms of geographic analysis, what role they play and how they are manifested in any particular analysis are less clear. Geography may still need what Harvey called for in *Explanation in Geography*: more transparency and clarity as to how laws, drawn from any discipline, integrate into the logical explanations posed by geographic theories. Goodchild (2004) suggests that in many applications the distinctions between laws, hypotheses, and theories may be unimportant because (1) they are readily substituted in common usage and (2) the true value of empirically valid, general statements comes from their simplicity and usefulness in application and prediction. While we agree with the later point, we nonetheless value distinguishing between theories, laws, and hypotheses when considering how replication can be used to build the evidence to support new geographic laws, or when examining how existing geographic laws complicate that process. If we understand laws as descriptive statements that are component parts of larger theoretical structures responsible for explanation, then we need not be concerned with the critique that existing geographic laws do not explain relationships or shed light on process. Distinguishing laws from theories, we can think of a law as a finely-tuned machine that is able to accomplish a task, but is also entirely ignorant of why it works (LaBracio, 2016). Explaining why a law works is the role of theory.

Through replication, we can attempt to determine the contexts in which a law does or does not function. If we are able to link candidate laws with theories that explain how some geographic phenomena operates, we will be more apt to design appropriate replication studies saving both resources and confusion. If our theories are correct, operating in the geographic contexts under examination, and are not confounded by localized conditions, replications will provide confirmatory evidence of our hypothesized relationships. If our theories are correct, but are not operating in the geographic context we are examining because of some local confound, replications can still provide evidential value. Observing a failed replication should motivate us to search for the reason for that failure. Hopefully, leading to the identification of necessary ceteris paribus conditions. Crucially, replications can provide this value even if we are not aware what the confounding conditions might be. The complication, and the task for geographers, is the need to distinguish between a poor theory and a good theory that is simply being blocked by some unknown local factors. The second law of geography tells us that we should expect condition to vary in space, but it does not tell us how those conditions might affect the explanations we are testing. Only the theory itself can provide that reasoning, and replications completed across locations provide evidence we can use to revise the logical structure of our theories.

While we focus our attention on Tobler’s law and Anselin’s heterogeneity principle, there are many other candidates for geographic laws that may well shape the use of replication in the discipline and our pursuit of theory. A detailed consideration of those laws is also warranted. As an example, Zhang and Goodchild’s (2002) principle that it is impossible to measure location or describe geographic phenomena exactly reinforces the fundamental need to identify sources of uncertainty and estimate their potential effects when using replication in geographic research. Perhaps more provocatively, the fractal principle (Goodchild and Mark, 1987) suggests that examinations at progressively finer spatial resolutions reveal more detail about geographic phenomena at a predictable rate. The fractal principle speaks directly to the amount of information we have about a system and forces us to consider how we measure proximity and variability if we wish to test potential laws when a replication examines phenomenon at different spatial resolutions. If a change in extent accompanies a change in resolution, then a rise in detail would need to be balanced against the variability of the systems and a potential change in the theoretical structure linked to a law. Whether the fractal principle holds for the uncertainty principle would likewise radically alter our approach to replication and the search for laws. If uncertainty changes predictably with resolution, we would need to recalibrate our assessments of replication with scale, but we might also be able to estimate what those calibrations should be based on the magnitude of rescaling.

We have already considered the ramification of two types of law specifications: deterministic and probabilistic laws. However, other typologies of geographic law may also provide scaffolding for important insight into designing replication studies and determining thresholds of empirical evidence. Golledge and

Amedeo's (1968) reformulation of Bergman (1958) into five different types of laws developed by geographers may be one fruitful starting point, as would Sack's (1973) distinction between laws with and without explicit spatial reference. For example, it would be useful to determine if different consistency criteria are needed when comparing a replication and an original study for a cross-sectional law or an equilibrium law. Whereas a cross-sectional law poses a functional connection between variables and can be assessed by comparing the magnitude and uncertainty of the variables across studies, an equilibrium law states that a change will occur if some conditions are met. To assess the similarity of studies testing an equilibrium law, consistency criteria are needed for the conditions and the change. Because an equilibrium law does not say what will occur if conditions are not met, it is essential to clearly and precisely identify the conditions to ensure that a replication is in fact analyzing the same relation.

The wide variety of research topics and research approaches that characterize the geographic sciences suggest that geographers will be best served drawing lessons about replicability and the search for laws from different fields. As Castree (2005) points out, there is no clear reason why the experimental sciences should be the only model for geographic research. The statement holds for computational research as well. Disciplines currently at the forefront of reproducibility and replicability research have focused on computational and experimental research (NASEM, 2019), capturing only a portion of the diverse methodological toolbox geographers and geographic information scientists use to understand the world. Geographic researchers should cast a wide net when searching for lessons pertinent to the use of replication in the discipline. For example, it would be natural for GIScientists interested in discovering laws of spatial reasoning or the interpretation of spatial information to look to cognitive psychology and the biomedical fields for innovations in experimental design and the cataloging and sharing of confidential participant information. Establishing a repository of eye tracking data or functional Magnetic Resonance Images similar to the OpenfMRI database (Poldrack et al., 2013) and Brain Images of Normal Subjects image-bank (Job et al., 2017) could help establish the cognitive baselines, pattern expectations, and pool of evidence needed as references in replication studies, which would in turn facilitate the search for credible laws. Similarly, human geographers who typically conduct observational research could draw lessons from the social sciences. For example, geographers searching for laws and processes shaping uneven patterns of spatial development could draw practices from economics where the prohibitive cost of data collection and the pace of economic change often make the recollection of data and replication impractical. Each of these efforts could be amplified by the geospatial software standards (Wang, 2010; 2016; OGC, 2020) and infrastructure (Richardson 2019) already being developed by those working in the segments of the discipline more reliant on computation. As the science behind the systems, GIScience could act as the bridge between segments of the discipline developing new approaches to replication.

Conclusion

Reviews of prominent geographic laws (Goodchild, 2004; Waters, 2016; Anseling & Li, 2020), careful treatments of the varied practices used in geographic research (Castree, 2005; Kitchin, 2006), and interrogations of the roots of positivist geography (Barnes, 2004; 2006; 2018; Sui & Kedron, 2021) have brought the question of laws into the current century. At the same time, provocative suggestions that data-driven geography may help resolve the discipline's nomothetic and ideographic dichotomy (see Miller & Goodchild, 2015) link the issue of laws to methodological trends at the forefront of the discipline's research agenda. This paper has sought to tie the search of geographic laws to the reproducibility of scientific research and the role of replication in scientific explanation. Geographers have yet to interrogate the mechanisms though which replication studies may accrue evidence for new laws of geography. And yet, our discussion here suggests that the first two laws of geography are likely to confound replication studies and law-seeking across the social and environmental sciences. This implies an urgent need for geographers to develop the conceptual underpinnings of replication in the discipline while also creating infrastructure and research protocols for replication studies in the context of spatial dependence and heterogeneity. As replication studies become increasingly possible with the development of infrastructure for more reproducible research, geographers will need to develop a series of studies to test how the laws of geography confound replication, and a set of shared principles for replication assessment based on fundamental laws of geography.

Although we have discussed replication and the formulation of geographic laws in the context of the first two laws of geography in this paper, the arguments presented here should be pursued further. The first two laws of geography are essentially statements about stability and interconnection in space. However, stability and interconnection can and should be addressed across time, scales, and within linked and nested systems. Doing so opens the question of replication far wider and invites the search for spatio-temporal laws that may advance our ability to explain geographic phenomena. Beyond geography, many forms of research in the social and natural sciences occur within a geographic context, making the first two laws of geography likely confounds of replication and generalizability across the sciences. The laws of geography therefore have implications for any other forms of disciplinary, human-environment, and convergence research seeking to build credible theory with empirical replications in geographic space.

References

Ackerman, E. A. (1945). Geographic training, wartime research, and immediate professional objectives. *Annals of the Association of American Geographers*, *35*(4), 121-143.

Allison, D. B., Shiffrin, R. M., & Stodden, V. (2018). Reproducibility of research: Issues and proposed remedies. *Proceedings of the National Academy of Sciences*, *115*(11), 2561-2562.

Anselin, L. (1989). What is special about spatial data? Alternative perspectives on spatial data analysis. Technical Report 89-4. Santa Barbara, CA: National Center for Geographic Information and Analysis.

Anselin, L., & Li, X. (2020). Tobler’s Law in a Multivariate World. *Geographical Analysis*. https://doi.org/10.1111/gean.12237

Barnes, T. J. (2004). Placing ideas: genius loci, heterotopia and geography's quantitative revolution. *Progress in Human Geography*, *28*(5), 565-595.

Barnes, T. J. (2006). Geographical intelligence: American geographers and research and analysis in the Office of Strategic Services 1941–1945. *Journal of Historical Geography*, *32*(1), 149-168.

Barnes, T. J. (2018). A marginal man and his central contributions: The creative spaces of William ('Wild Bill') Bunge and American geography. *Environment and Planning A: Economy and Space*, *50*(8), 1697-1715.

Bergmann, G. (1957). *Philosophy of Science*. University of Wisconsin Press: Madison.

Bergson, H. (1950). *An Introduction to Metaphysics*. The Liberal Arts Press: New York.

Braithwaite, R. B. (1960). *Scientific Explanation*. Harper Torchbooks: New York.

Brunsdon, C. (2016). Quantitative methods I: Reproducible research and quantitative geography. *Progress in Human Geography*, *40*(5), 687-696.

Brunsdon, C., & Comber, A. (2020). Opening practice: supporting reproducibility and critical spatial data science. *Journal of Geographical Systems*, 1-20.

Bunge, W. (1966), *Theoretical Geography* (2nd ed.). Lund Studies in Geography. CWK Gleerup.

Castree, N. (2005). Is geography a science. *Questioning Geography. Oxford: Blackwell Publishing*, 57-79.

Chorley, R., Haggett, P. (eds) (1967) *Models in Geography*. Methuen, London.

Coffman, L. C., and Niederle, M. (2015). Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives* 29 (3), 81–98.

Davies, W. K. (1968). The need for replication in human geography some central place examples. *Tijdschrift voor economische en sociale geografie*, *59*(3), 145-155.

Earp, B. D., and D. Trafimow. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology* 6, 621.

Golledge, R., & Amedeo, D. (1968). On laws in geography. *Annals of the Association of American Geographers*, *58*(4), 760-774.

Gregory, D. (1978) *Ideology, Science and Human Geography*. Hutchinson, London.

Goodchild, M. F. (2004). The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*, *94*(2), 300-303.

Goodchild, M. F., & Mark, D. M. (1987). The fractal nature of geographic phenomena. *Annals of the Association of American Geographers*, *77*(2), 265-278.

Goodchild, M. F., Fotheringham, A. S., Kedron, P., & Li, W. (2020). Introduction: Forum on reproducibility and replicability in geography. *Annals of the American Association of Geographers*, 1-4.

Guelke, L. (1971). Problems of scientific explanation in geography. *Canadian Geographer/Le Géographe canadien*, *15*(1), 38-53.

Guelke, L. (1977). The role of laws in human geography. *Progress in human Geography*, *1*(3), 376-386.

Guelke, L. (1978) Geography and logical positivism. In Johnston, R. and Herbert, D. (eds) *Geography and the Urban Environment*, vol. 1. Wiley, Chichester, pp. 35–61.

Haggett, P., Chorley, R. J., & Stoddart, D. R. (1965). Scale standards in geographical research: a new measure of areal magnitude. *Nature*, *205*(4974), 844-847.

Hartshorne, R. (1939). The nature of geography: A critical survey of current thought in the light of the past. *Annals of the Association of American Geographers*, *29*(3), 173-412.

Hartshorne, R. (1954). Comment on" Exceptionalism in geography". *Annals of the Association of American Geographers*, *44*(1), 108-109.

Hartshorne, R. (1955). “Exceptionalism in geography’’ Re-examined. *Annals of the Association of American Geographers*, *45*(3), 205-244.

Harvey, D. (1969). *Explanation in Geography*. Edward Arnold Press.

Hay, A. (1979) Positivism in human geography: response to critics. In Johnston, R.J. and Herbert, D. (eds) *Geography and the Urban Environment*, vol. 2. Wiley, Chichester, pp. 1–26.

Hay, A. (1985) Scientific method in geography. In Johnston, R.J. (ed.) *The Future of Geography*. Methuen, London, pp. 129–142.

Hemple, C. G. (1965). *Aspects of Scientific Explanation.* MacMillan: New York.

Job, D.E., et al. (2017) A brain imaging repository of normal structural MRI across the life course: Brain Images of Normal Subjects (BRAINS). *NeuroImage* 144: 299-304.

Jones, Emrys (1956). "Cause and effect in human geography". *Annals of the Association of American Geographers*. **46:** 369-77

Kedron, P., Frazier, A. E., Trgovac, A. B., Nelson, T., & Fotheringham, A. S. (2019). Reproducibility and replicability in geographical analysis. *Geographical Analysis*.

Kedron, P., Li, W., Fotheringham, S., & Goodchild, M. (2020). Reproducibility and replicability: opportunities and challenges for geospatial research. *International Journal of Geographical Information Science*, 1-19.

Kincaid, H. 1990. Defending Laws in the Social Sciences. *Philosophy of the Social Sciences* 20 (1):56–83.

Kitchin, R. (2006). Positivistic geographies and spatial science. In S. Aitken, & G. Valentine (Eds.), *Approaches to human geography* (pp. 20-29). SAGE Publications Ltd, https://www.doi.org/10.4135/9781446215432.n2

Konkol, M., Kray, C., & Pfeiffer, M. (2019). Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. *International Journal of Geographical Information Science*, *33*(2), 408-429.

LaBracio, L. (2016, June 7). What’s the difference between a scientific law and theory? *Ted-Ed Blog.* https://blog.ed.ted.com/2016/06/07/whats-the-difference-between-a-scientific-law-and-theory-in-ted-ed-gifs/

Latour, B. (2005). Reassembling the Social: An Introduction to Actor-Network-Theory. Oxford University Press.

Lewis, P. W. (1965). Three related problems in the formulation of laws in geography. *The Professional Geographer*, *17*(5), 24-27.

Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. *GeoJournal*, *80*(4), 449-461.

Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace and World: New York.

National Academies of Sciences, Engineering, and Medicine (NASEM), 2019. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25303.

Nichols, J. D., M. K. Oli, W. L. Kendall, & G. S. Boomer. (2021). A better approach for dealing with reproducibility and replicability in science. *Proceedings of the National Academy of Sciences of the United States of America* 118 (7):1–5.

Nüst, D., C. Granell, B. Hofer, M. Konkol, F. O. Ostermann, R. Sileryte, Cerutti, V. (2018). Reproducible research and GIScience: an evaluation using AGILE conference papers. *PeerJ* 6:e5072.

Nüst, D., Ostermann, F. O., Granell, C., & Kmoch, A. (2020). Improving reproducibility of geospatial conference papers–lessons learned from a first implementation of reproducibility reviews. Septentrio Conference Series, No 4 (2020): The 15th Munin Conference on Scholarly Publishing 2020. DOI: https://doi.org/10.7557/5.5601

Nüst, D., and E. Pebesma. 2021. Practical Reproducibility in Geography and Geosciences. *Annals of the American Association of Geographers* 111 (5):1300–1310.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science* 349 (6251):aac4716. http://science.sciencemag.org/content/349/6251/aac4716.abstract.

Peet, R. (1999) *Modern Geographical Thought*. Blackwell, Oxford.

Poldrack, R.A., et al. (2013) Toward open sharing of task-based fMRI data: the OpenfMRI project. *Frontiers in Neuroinformatics* 7: 12.

Popper, K. (1959) *The Logic of Scientific Discovery*. Hutchinson: London.

Rey, S. J. (2009). Show me the code: spatial analysis and open source. *Journal of Geographical Systems* 11 (2):191–207.

Richardson, D., (2019). Dealing with Geoprivacy and Confidential Geospatial Data Crossing Borders. *ARC News*, *41*(2).

Roesch, E., & Rougier, N. (2020). New journal for reproduction and replication results. *Nature*, *581*(7806), 30-30.

Rose, J. A. (1966). Dissent from Down under: Metropolitan Primacy as the normal State*. *Pacific Viewpoint* 7 (1):1–27.

Sack, R. D. (1972). Geography, geometry, and explanation. *Annals of the Association of American Geographers*, *62*(1), 61-78.

Sack, R. D. (1973). A concept of physical space in geography. *Geographical Analysis*, *5*(1), 16-34.

Sack, R. D. (1980). *Conceptions of space in social thought*. Macmillan International Higher Education.

Sayer, A. (1992). *Method in Social Science* (2nd ed). Routledge: London

Schaefer, F. K. (1953). Exceptionalism in geography: A methodological examination. *Annals of the Association of American geographers*, *43*(3), 226-249.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of general psychology*, *13*(2), 90-100.

Searle, J. (1984). *Minds, brains and behavior.* Cambridge, MA: Harvard University Press.

Peter Sjøholt. 1984. Some problem areas in transferring central place theory to developing countries, *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, 38:1, 27-35.

Stigell, E., & Schantz, P. (2011). Methods for determining route distances in active commuting–Their validity and reproducibility. *Journal of Transport Geography*, *19*(4), 563-574.

Sui, D., and P. Kedron. 2021. Reproducibility and Replicability in the Context of the Contested Identities of Geography. *Annals of the American Association of Geographers* 111 (5):1275–1283.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, *46*(sup1), 234-240.

Tullis, J. A., & Kar, B. (2021). Where Is the Provenance? Ethical Replicability and Reproducibility in GIScience and Its Critical Applications. *Annals of the American Association of Geographers* 111 (5):1318–1328.

Wainwright, J. (2020). Is critical human geography research replicable? *Annals of the American Association of Geographers*, 1-7.

Wilson, J. P., K. Butler, S. Gao, Y. Hu, W. Li, and D. J. Wright. 2021. A Five-Star Guide for Achieving Replicability and Reproducibility When Working with GIS Software and Algorithms. *Annals of the American Association of Geographers* 111 (5):1311–1317.

Wang, S. (2010). A CyberGIS framework for the synthesis of cyberinfrastructure, GIS, and spatial analysis. *Annals of the Association of American Geographers*, *100*(3), 535-557.

Wang, S. (2016). CyberGIS and spatial data science. *GeoJournal*, *81*(6), 965-968.

Waters, N. (2016). T obler's First Law of Geography. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, 1-15.

Waters, N. (2020). Motivations and methods for replication in geography: Working with data streams. *Annals of the American Association of Geographers*, 1-9.

Zhang, J. X., Goodchild, M. F. (2002). *Uncertainty in geographical information*. Taylor and Francis: New York.