- Poisson Permutation Test to Locate Transition from Extensive to Intensive Development

Growth in the number of scholars and the number of published scholarly works is attended by a qualitative transitions between extensive and intensive patterns of citation. When disciplines are very young scholars are almost always exploring new or at least unclaimed terrain with little interest in covering the same ground twice. As disciplines develop a transition invariably occurs; scholars become much more likely to retrace familiar ground. Much of the work of this study aims to understand the significance of this fact.

First the fact must be established. Intellectual terrain is often imagined as a space of meaning or a population meaningfully distinguishable ideas. Because ideas cannot be directly observed, several indicators of their presence have been used, citations chief among them. Empirically, then, we will start on the basis of the citations as a useful indicator of ideas. Later we will discuss the limitations of the ideational theory, and we will present an alternative interpretation of a citation space. Luckily the facts at issue will not change.

So, then, it will be useful to treat the terrain of scholarship as the accumulating stock of already cited references. The act of exploration, so to speak, of this space consists in the inclusion, in the reference list of a scholarly publication, of a particular set of citations and not another. A footprint in this space, left by one publication, may be represented as a count of each citation pair in the list of references. This operationalization allows footprints to overlap completely, partially, or not at all. By enumerating citation pairs or co-citations instead of their individual counts, we also claim that the meaning of a reference may vary in combination with other references.

A more empiricist and less theory laden interpretation is to claim that we may identify how disparate acts of cultural production hang together, without knowing why they do so. Citations provide merely one kind of thread, but were we to trace out several more modes of relatedness then we might provide a fuller picture of the sociocultural structure underpinning scholarship. Such a task is beyond scope for the present study, but we can at least specify ignorance ({**???**}). Clearly there is much more to the content of a publication than its list of references. But even considering this narrow slice of its meaningful content, we are already at pains to generalize from the observation of a citation pattern to the cause of that pattern appearing in a particular time and place. It will be difficult, for instance, to posit a choice mechanism, for we cannot discern whether the inclusion of a reference was the choice of the author, the editor of a journal, the reviewers refereeing the manuscript, a colleague listed in the acknowledgements, an uncredited inspiration, etc. I therefore make no effort to identify an actor responsible for an included reference, but rather consider it the outcome of the local art world surrounding the production of that piece of scholarship {[}c.f. theories of authorship @ ;@ {]}. What is a critical problem to solve for the intellectual historian may a fool's errand for the population researcher. It is a mistake to treat any particular citation, and especially to

treat the entire reference list, as reflective of the choice of the author. Indeed this mystifies the production process behind scholarship.

An extensive pattern of citation then is one that both introduces never before cited references and one that favors those extant references that have been cited the least by others. A Poisson distribution is a simple first approximation of a random search in this space, and observed citation counts with a mean below the random pattern (underdispersion) can be considered to represent the extensive pattern, while means above the same (overdispersion) may represent the intensive pattern.

This extensive pattern of development may be compared to the paradigmatic model described by Kuhn ({**???**}10). Once a paradigm, in the sense of a model to be extended, takes hold among a community of scholars, normal science ensues as a process of narrowing the range of possibilities opened by the paradigm. The specifically scientific pattern of development is to retrace familiar problems until they are solved, and then to relegate the solution to one or another form of black box, such as mathematical codification, textbook explication, or codification in technology. Familiar ground is in one moment intensely retraced, and in the next systematically forgotten. Indeed Kuhn aims to demonstrate that the ideology of cumulative development in the sciences is a consequence of black boxing, which serves to render subsequent generations of scientists ignorant of a history better described by a cyclical or sinusoidal trend.

While pre-history of disciplines are beyond Kuhn's scope, here they are paramount. This emphasis is based in a hunch that the mechanisms that govern the genesis of disciplines may be implicated in their ongoing development.

References