Fall 9-29-2011

Heterogeneity and data analysis

Peter J. Taylor
University of Massachusetts Boston, peter.taylor@umb.edu

WORKING PAPER on Science in a Changing World

Paper # 8 -2011

Heterogeneity and data analysis

PETER J. TAYLOR

GRADUATE PROGRAM IN CRITICAL AND CREATIVE THINKING
University of Massachusetts, Boston, MA 02125, USA
http://scholarworks.umb.edu/cct_sicw/8

Heterogeneity and Data Analysis

Peter J. Taylor
Graduate track in Science in a Changing World, http://sicw.wikispaces.umb.edu
University of Massachusetts, Boston, MA 02125, USA. peter.taylor@umb.edu

Abstract

This working paper is a discussion paper for a September 2011 meeting of the research group of Prof. Di Cook on data visualization and exploratory data analysis at Iowa State University. A taxonomy of eleven kinds of heterogeneity is presented, followed by a set of vignettes that illustrate some of the meanings and sketch some implications, then a series of images that illustrate the heterogeneities. Several of the vignettes speak to a broad contention about heterogeneity and control: In relation to modern understandings of heredity and development over the life course, research and application of resulting knowledge are untroubled by heterogeneity to the extent that populations are well controlled.

Preamble

This working paper is a discussion paper for a September 2011 meeting of the research group of Prof. Di Cook on data visualization and exploratory data analysis at Iowa State University.[1] My interest in visual exploration of large data sets traces back to my studies and first research job in multivariate "pattern analysis" in ecology and agriculture (in Australia in mid 1970s). Conversations among the plant breeders I worked with[2] were lively when they saw the plots I generated for them. Much less so when I showed them analyses of variance and other numerical output. Although I have since strayed from my quantitative roots—I am now more of a sociologist and philosopher of science than a data analyst—I remain very interested in ways that people push the limits of conventional quantitative methods.

The theme of people addressing or suppressing *heterogeneity* runs through my studies these days of what researchers do (or don't do) in social epidemiology, population health, and quantitative genetics. In this vein, *I see the various tools of interactive and dynamic graphics for data analysis as ways to address heterogeneity, in the sense of teasing apart homogeneous components of a (heterogeneous) mixture so that separate kinds of explanations can be formulated for the separate components*. Traditional statistical analysis allows itself to be confounded by the mixture of patterns or structure in a given data set. In this spirit, Cook and Swayne (2007, 13) quote Buja (1996) approvingly: “Non-discovery is the failure to identify

[1] In the session, I gave a brief introduction, then asked participants to take turns, say 5 minutes each, to relate how the paper intersected with or stimulates their own thinking (while the author stayed quiet,
[2] The P.I. for this first research job, Don Byth, was a 1965 Iowa State Ph.D.

meaningful structure… [T]he fear of non-discovery should be at least as great as the fear of false discovery."

Our discussion may shed light on why the issue of heterogeneity is not explicitly named in discussions of exploratory data analysis by interactive graphics. If my characterization of this enterprise makes sense to you, how does your experience with exploratory data analysis and interactive graphics helps you think about the range of meanings of heterogeneity? *This paper by no means circumscribes the issues you might bring to the topic of heterogeneity and data analysis. Nor do I presume that the vignettes resonate with your day-to-day concerns.* Yet, I do hope some of these thoughts-from-an-outsider stimulate discussion in which specialists in representation and analysis of data provide deeper accounts of the conceptual and practical issues, correct my presentation when necessary, and help me learn more.

Section I presents my current taxonomy of heterogeneities. The vignettes in Section II illustrate some of the meanings and sketch some implications.[3] Section III presents images, not all self-explanatory, that illustrate the taxonomy. Several of the vignettes speak to a broad contention presented in a Coda concerning heterogeneity and control.

[3] These vignettes are extracted or adapted from publications, blogposts, notes to students, unpublished drafts, and a proposal for a book on heterogeneity in the biomedical sciences.

I. A Taxonomy of Heterogeneities

Static	1. There is an assortment, each a separate type ("cabinet of curiosities")
	2. Mixture of types (e.g., allelic heterogeneity & locus heterogeneity in genetics)
	3. Trait = composite of types (analogy: the 3 components of a triathalon)
Variational	4. There is noise or error, but that is deviation from the type or essential trajectory
	5. Variation in a set of traits involves a composite of variance/covariance structures (statistical heterogeneity)
	6. There is variation, not types
	7. When similar responses of different individual (e.g., genetic) types are observed, it is not necessarily the case that similar conjunctions of risk or protective factors have been involved in producing those responses (=possibility of "underlying heterogeneity")
Dynamic	8. Variation produces qualitative changes in results from standard theory based on uniform units (e.g., theory about Malthusian population growth, tragedy of the commons, prisoner's dilemma)
	9. Heterogeneity in pathways of development Variants from Taylor (2005): 9a. "Intersecting processes" Processes operating at different spatial and temporal scales that cut across the boundaries of the situation under consideration and restructure its "internal" dynamics. 9b. "Unruly complexity," which arises whenever there is ongoing change in the structure of situations that have built up over time from heterogeneous components and are embedded or situated within wider dynamics. 9c. “Heterogeneous construction,” in which, researchers establish knowledge and technological reliability through practices that are developed through diverse and often modest practical choices. This is the same as saying the researchers are involved in contingent and on-going mobilizing of diverse materials, tools, people, and other resources into webs of interconnected resources.
Dynamic-participatory	10. Participatory restructuring of the dynamics (intersecting processes, unruly complexity, or heterogeneous construction) that generated the data.
	11. Participatory restructuring through multiple points of engagement, which occurs in tension with deployment or withholding of trans-local knowledge and resources.

II. Vignettes

Heterogeneity #1, An assortment.

In an essay on "The Analytical Language of John Wilkins," Borges (1964) mentions a “doctor Franz Kuhn” referring “to a certain Chinese encyclopaedia entitled 'Celestial Empire of benevolent Knowledge'. In its remote pages it is written that the animals are divided into: (a) belonging to the emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) et cetera, (m) having just broken the water pitcher, (n) that from a long way off look like flies.” (http://alamut.com/subj/artiface/language/johnWilkins.html. Michel Foucault brought attention to this passage in Borges in the opening to his The Order of Things.)

Heterogeneity #2, Mixture of types

If the population in question contains a mix of different types that are identified and separable, then identification of a subject’s type allows them to be treated or investigated separately and differently.

• Phenylketonuria (PKU) is a condition associated with a single genetic locus, but there are different mutations within that locus. People with PKU are a mix of people with different mutations or genetic sub-types.

• "Maternal PKU" arises when a fetus gestates in high-phenylalanine conditions that occur when a PKU mother was not insufficiently compliant with the diet (where compliance may be influenced by a variety of factors). If maternal PKU is considered a form of PKU, then the population of children with the deleterious symptoms of PKU is a mix of those with a genetic condition that was not followed by the special diet and those without the genetic condition who have.

• A certain population of light-eyed, yellow rats consisted of two strains, each bred separately from some ancestral founding group. The "two strains of light-eyed, yellow rats, each of which bred true by itself... produced nothing but black-eyed rats when crossed with each other" (Wright 1920, 37). (If each strain had been bred in the same, uniform laboratory conditions, this would seem to be a case of different kinds of genetic factors producing light-eyes for the two strains in those conditions.)

• The protective effect with respect to heart disease and stroke of taking a daily low dose of aspirin differs on average for men and women. This means the human population can be treated as a mix of female and male types with respect to the protective effects of aspirin. However, 9% of the patients in one study appeared "resistant," i.e., their samples didn't show the typical blood-thinning effects of aspirin, meaning that the male and female types are heterogeneous with respect to resistance. In practice, these sub-

types are costly to identify and this is rarely done (Eikelboom 2003). Note: It is also possible that the heterogeneous factors underlying the sub-types overlap, that is, the male and female types (which were based on differences on average) can eventually be resolved into a number of types not unique to males or females.

Heterogeneity #2, Mixture of types -> #7, Possibility of "underlying heterogeneity," and vice versa.

• In genetics, homogeneity may be on the surface only, e.g., when it is discovered that different genetic conditions are expressed as the "same" clinical entity. Conversely, the clinical expression of mutations at a single genetic locus can vary significantly from one person to the next (Kaplan 2000, 18). This may be because the mutations are at different points within the locus or because the same genetic condition develops in different genetic and environmental contexts, i.e., the other genetic and environmental factors vary among the people.

• In medical sociology Brown and Harris (1989) often find common meaning among subjects' different types of experience. In other words, Brown and Harris code sameness despite surface heterogeneity.

Heterogeneity #2, Mixture of types

Evaluations of Closed Circuit television (CCTV) as described by Tilley (2000) might be subject to a meta-analysis. However, as Tilley's lists below indicate, such an analysis would mix together studies of situations in which different mechanisms (or a mix of mechanisms) and different contexts apply. What meaningful recommendation could emerge form the meta-analysis, even if all results were in the same direction?

I was asked by officials at the Home Office to look at the effectiveness of the introduction of close circuit television in car parks as part of the Safer Cities Programme which was aiming to deal with local crime problems in 20 cities in England… First I thought about mechanisms. How might close circuit television affect rates of car crime?

Here is a list of mechanisms:

a) The 'caught in the act' mechanism. CCTV might reduce car crime by increasing the chances that current offenders are seen on screen detected committing their crimes and arrested, taken away, punished and deterred.

b) The 'you've been framed' mechanism. CCTV might reduce car crime by leading potential offenders to avoid the perceived risk that they might be caught and convicted because of the evidence on tape.

c) The 'nosy parker' mechanism. CCTV might lead to increased usage of car parks since drivers feel more safe. Their increased usage might then increase natural surveillance deterring potential offenders worried that they might be seen committing their crimes.

d) The ‘effective deployment’ mechanism. CCTV might enable security staff to be deployed more quickly where suspicious behaviour was going on. They then act as visible guardians.

e) The ‘publicity’ mechanism. CCTV and signs announcing its installation might symbolise efforts to take crime seriously and to reduce it. Potential offenders might want to avoid the perceived increased risk.

f) The ‘time for crime’ mechanism. Offenders might calculate that car crimes taking a long time risk their being caught on camera and they might decide only to commit those car crimes that could be completed very quickly.

g) The ‘memory jogging’ mechanism. The presence of CCTV and associated notices may remind drivers that their cars are vulnerable and lead them to lock them and operate security devices and remove easily stolen items from view.

h) The ‘appeal to the cautious’ mechanism. Cautious drivers sensitive to the possibility that their cars may be vulnerable to crime may use car parks with more security devices and displace less cautious drivers to other car parks. The high level of security of the car park users may make it difficult for offenders successfully to commit their crimes.

Having thought about mechanisms I then thought about context. Are all car parks and all car park crime problems the same? Well, here are some of the variations that I identified.

1. The ‘criminal clustering’ context. A given rate of car crime may result from a small number of very active offenders or a large number of occasional offenders. A mechanism leading to the offender being disabled holds promise according to the offender/offence ratio as in (a) above.

2. The ‘style of usage’ context. Long stay car parks fill up early in the morning and empty after work in the evening. If the dominant CCTV mechanism turns out to be increased confidence and usage, as in (c) or (h) above, then this will have little impact because the pattern of usage is already high, with little movement dictated by working hours not fear of crime. If, however, the car park is little used, but has a very high per user car crime rate, then increased usage mechanisms may lead to an overall increase in the number of crimes but a decreased rate per use.

3. The ‘lie of the land’ context. Cars parked in CCTV blind spots will be more vulnerable if the mechanism is increased chances of apprehension through evidence on video tape as in (b), but not if it is through changed attributes or security behaviour of customers, as in (g) or (h).

4. The ‘alternative targets’ context. The local patterns of motivation of offenders, together with the availability of substitute targets, provide the context for potential displacement elsewhere.

5. The ‘resources’ context. In isolated car parks with no security presence and no police near to hand the deployment of security staff or police as a deterrent as in (d) is not possible.

This is not, of course, necessarily a comprehensive list of contexts or mechanisms. What it brings out, though, is that even in relation to a relatively simple measure in a relatively simple setting the range of mechanisms and contexts is quite wide. It is unlikely that closed circuit television will have the same effect

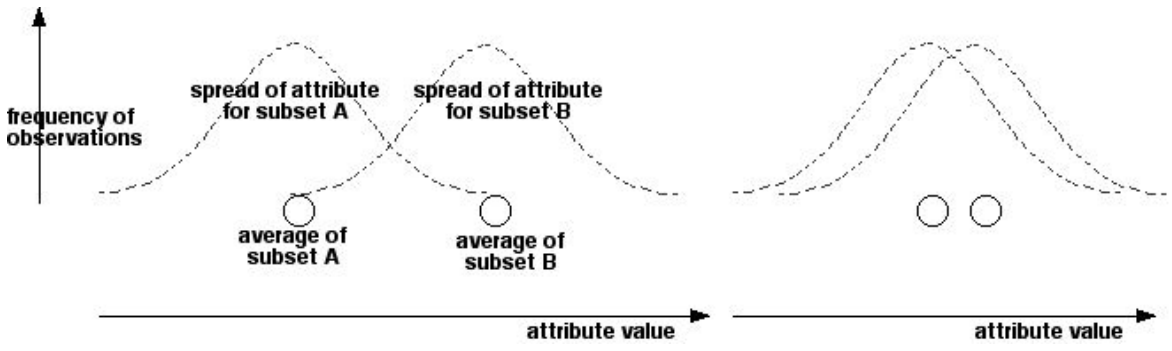
on car crime rates in all circumstances. The mechanisms and contexts are just too varied. Added to this, of course, CCTV itself varies substantially in its technical capacity, which will affect its potential to trigger some of the mechanisms which have been identified here. The issue for the evaluator is that of working out how to test, or arbitrate between, a variety of theories that explain how and where CCTV might have its impact on car crime.

Heterogeneity #4, Deviation from the type

Statistical analysis rests on the simplest heterogeneity, namely, variation around a mean. In this light, I tell education students who will not be taking a statistics course that they should:

Understand the simple chain of thinking below, then enlist or hire a statistician who will use the appropriate recipe for the data at hand.

1. There is a population of individuals. (Population = individuals subject to the same causes of interest. In addition to these foreground causes, there may also be background, non-manipulatable causes that vary among these individuals.)
2. Variation: For some measurable attribute, the individuals have varying responses to these causes (possibly because of the background causes).
3. You have observations of the measurable attribute for two or more subsets (samples) of the populations.
4. Central question of statistical analysis: Are the subsets sufficiently different in their varying responses that you doubt that they are from the one population (i.e., you doubt that they are subject to all the same foreground causes)? Statisticians answer this question with recipes that are variants of a comparison between the subset averages in relation to the spread around the averages. For the figure below, the statisticians' comparison means that you are more likely to doubt that subsets A and B are from the same population in the left hand situation than in the right hand one.



5. If you doubt that the subsets are from the same population, investigate further, drawing on other knowledge about the subsets. You hope to expose the causes involved and then take action informed by that knowledge about the cause.

Variation around a mean is not a strong sense of heterogeneity. The emphasis above is on the means (the circles) more than the variation (the dashed curves). Statistical analysis distinguishes types (or decides they are not distinguishable) more than it explores the variation (or error, i.e., deviation from type). Data amenable to a t-test are, however, open to alternative explorations, as illustrated by the final vignette in this discussion paper.

Heterogeneity #4, Deviation from the type or essential trajectory

Of course, statistical analysis involves more than t-tests and their generalizations. Correlation and regression are another mainstay. Here, however, the emphasis lies more on prediction than variation, as if, as a generalization of the emphasis in t-tests on types, the line or curve of prediction captured the *essential trajectory* of the data (McLaughlin 1989). (Of course, everyone knows that correlation is not causation, but most of us interpret regressions in a causal spirit.) The following excerpt from Taylor (2008; see http://bit.ly/osTjQ3) highlights an alternative view of correlation and regression that keeps our attention on the variation:

> Consider the concept of a regression line as a best predictor line. To predict one measurement from another is to hint at, or to invite, causal interpretation. Granted, if we have the additional information that the second measurement follows the first in time—as is the case for offspring and parental traits—a causal interpretation in the opposite direction is ruled out. But there is nothing about the association between correlated variables, whether temporally ordered or not, that requires it to be assessed in terms of how well the first *predicts* the second (let alone whether the predictions provide insight about the causal process). After all—although this is rarely made clear to statistics students—the correlation is not only the slope of the regression line when the two measurements are scaled to have equal spread, but it also measures how tightly the cloud of points is packed around the line of slope 1 (or slope -1 for a negative correlation). Technically, when both measurements are scaled to have a standard deviation of 1, the average of the squared perpendicular distance from the points to the line of slope 1 or -1 is equal to 1 minus the absolute value of the correlation (Weldon 2000). This means that the larger the correlation, the tighter the packing. This tightness-of-packing view of correlation affords no priority to one measurement over the other. Whereas the typical emphasis in statistical analysis on prediction often fosters causal thinking, a non-directional view of correlation reminds us that additional knowledge always has to be brought in if the patterns in data are used to support causal claims or hypotheses.

[Postscript: The tightness of packing view of regression for continuous variables can be extended to multivariate associations through Principal Component Analysis, factor analysis, etc. The well-known difficulty of interpreting principal components or the factors can be flipped on its head: What causal

assumptions about *independent* variables (i.e., independently modifiable variables) enter into interpretations of conventional regression analysis?]

Heterogeneity #4, Deviation from the type or essential trajectory -> Heterogeneity #6: Variation, not types -> Heterogeneity #9. Heterogeneity in pathways of development

Imagine a comparison of the dental health of two communities that have the same range of health problems except that the one with naturally high level of fluorides in its water supply has better than average dental health. In each community there will be variation around the average dental health. However, if the variation is small relative to the differences in the two averages, it might seem reasonable to advocate fluoridation of water supplies lacking natural fluoride. In doing so the variation around the average (the very simplest form of heterogeneity) is discounted, as are other deviations from type, such as teeth discoloration that occurs in some individuals. Public health policy-makers discount the variation because the benefits exceed the costs when summed up for the community. The policy-makers are able to do this as long as the infrastructure for water-supply fluoridation remains part of public expenditures covered by taxpayers and as long as individuals who bear disproportionate cost (e.g., those who teeth are discolored) do not effectively mobilize resources and allies to resist—in other words, as long as the population is well controlled. Opponents of fluoridation of the water supply who accept the data on benefits and costs (many opponents do not; Colquhoun 1997) could still promote a participatory alternative: fluoride tablets to be taken by each individual, which would allow people subject to teeth discoloration to adjust the dosage or to choose to manage their dental health without fluoride. This approach is not preferred by most public health policy-makers, who point to lack of "compliance" when individuals are responsible for administering their own preventative medicines. Participation is seen as unreliable; control is more effective. Population health is the guiding idea; variation within the communities is not troubling (Rose 2008).

Suppose now that two "racial" groups show persistent differences on average in some scholastic achievement tests (where racial categories are as defined, say, by the U.S. census). By analogy with the fluoride case, we should ascribe the difference to race, that is, to some social or biological variable(s) that differ from one race to the other. Identifying those variables will not be as simple as noting the presence or absence of fluoride, but should researchers even try to find them? What if they were to succeed?—If the variable were unalterable (say, a matter of genes), would we resign ourselves to the difference? If the variables were biologically or socially alterable, would we administer the same "antidote" to all in the lower-achieving group? What kind of social infrastructure would be involved? (Think here of No Child Left Behind measures mandated in the name of decreasing racial disparities in K-12 test results.)

In contrast to the fluoridation case, we can readily imagine researchers and policy-makers, unhappy with

explanations and policies based on group membership, who want to shift the focus to the heterogeneous pathways of development, in this case, of scholastic achievement. Given the social context in which such a move would be envisaged and enacted, these researchers and policy-makers are likely to face troubling tensions or conundrums—How can attention be given to diversity of pathways without bolstering the popular fiction that racial group membership in the United States no longer brings social benefits and costs and without providing support for various initiatives that have been attempting to prohibit the collection and use of racially classified information by state and local governments (e.g., the failed 2003 Proposition 54 in California)? At the same time, the racial categories used in censuses and other surveys continue to change—as does people's identification with those categories (Hirschman et al. 2000)—yet longituidinal analysis depends on data collected under the same categories for extended periods of time. In short, researchers and policy-makers concerned about heterogeneity within and across racial groups have to use data collected under racial categories, and, despite the shifting nature of those categories, get drawn into defending the continued collection of such data lest there be no information and thus no pattern (such as the average IQ test score differences) to push away from (Taylor 2009a).

It is always illuminating to reconstruct how researchers negotiate the tension between, on one hand, analyses and action based on averages for groups or populations, and, on the other hand, paying attention to variation from those averages and heterogeneous pathways of development.

Heterogeneity #7, Possibility of "underlying heterogeneity"

Different kinds or combinations of factors are involved in what is deemed the same response. The challenge is to expose the factors and the ways they contribute to the response in question, if that is possible.

• Consider the height a high jumper jumps. The athlete may use the classical approach to the jump and movements in the air or those of the Fosbury flop.

• Studies of heritability of human traits associate the similarity among twins or a set of close relatives with similarity of (yet-to-be-identified) genes or genetic factors. ("Heritability" is a technical term with a statistical basis, readily confused with, but quite distinct from, the colloquial idea of genes transmitted from parents to offspring; Taylor 2010a.) The methods of data analysis cannot rule out the possibility that the factors underlying the development of observed traits may be heterogeneous. That is, although relatives may be similar for a given trait because they share more genes or environmental conditions than unrelated individuals, the genes and environmental conditions underlying the development of the trait *need not be the same from one set of relatives to another*. As illustrated in Figure 1, it could be that pairs of genetic variants (alleles) at a number of positions on the genome, say, AAbbccDDee, subject to a sequence of environmental factors, say, FghiJ, are associated, all other things being equal, with the same

outcome for the trait as are variants aabbCCDDEE subject to a sequence of environmental factors FgHiJ (Taylor 2010a).

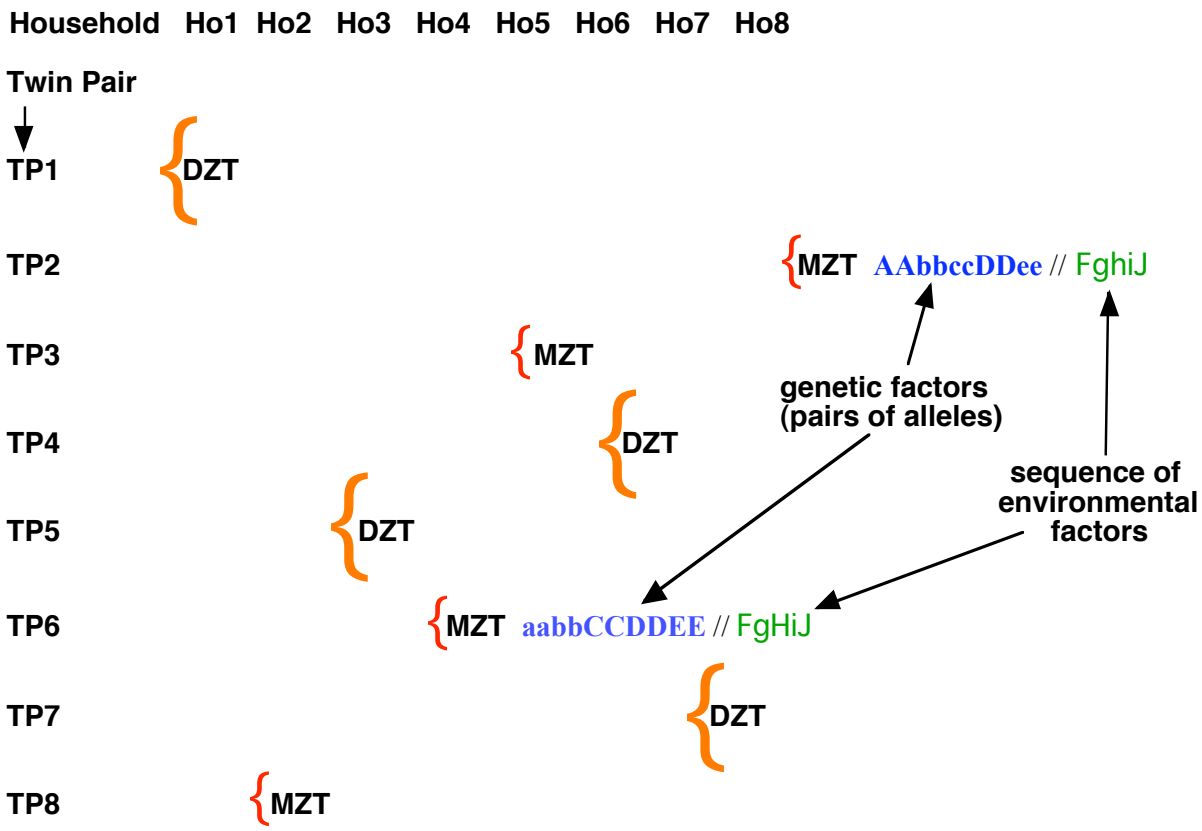


Figure 1. Factors underlying a trait may be heterogeneous even when identical (monozygotic) twins raised together (MZT) are more similar than fraternal (dizygotic) twins raised together (DZT). (The greater similarity is indicated by the *smaller* size of the curly brackets.) The underlying factors for two MZT pairs are indicated by upper and lower case letters for pairs of alleles (A-E) and environmental factors to which they are subject (F-J).

Some prominent geneticists have noted that heritability estimates are not helpful in identifying specific genetic factors (e.g., Rutter 2002, 4), but the possibility that the underlying genetic and environmental factors influencing development of a trait may be heterogeneous has yet to be recognized as a significant methodological concern by quantitative geneticists or by critical commentators on heritability research (e.g., Downes 2004 and references therein, but see Taylor 2008a). However, the common use of heritability as a basis for judging a trait to be a good candidate for molecular research (e.g., Nuffield Council on Bioethics 2002) is not so helpful if underlying factors can be heterogeneous (Taylor 2010a). In the case of agricultural breeding (where quantitative genetics originated) the absence of attention to

underlying heterogeneity can be understood given that researchers have enough control of their varieties and conditions in test locations to take compensatory steps when results of selection informed by heritability studies (and related data analysis) do not meet predictions. Moreover, the agricultural extension system allows recommendations to farmers that match varieties with conditions of cultivation or husbandry (Taylor 2009b). Such control over materials and human subjects (through an established social infrastructure for providing advice that shapes their practice) is not, however, readily available to social scientists and other commentators on the nature-nurture debate. Nor is it straightforward to control the subjects of human molecular biology and biotechnological advances. However, one may look for subclasses in which underlying factors are uniform, not heterogeneous. If found, use research to probe and treat this subclass. Attempts to extrapolate back to other subclasses are likely to follow. Will they be successful?

Two issues are raised by this vignette: the possible heterogeneity of factors that underlie observed traits warrants attention; and the lack of attention to it invites historical, sociological, and philosophical interpretation.

Heterogeneity #9, Heterogeneity in pathways of development -> potential for #11, Participatory restructuring through multiple points of engagement

> The man of the moment [was] J. Craig Venter, Ph.D., whose pioneering work to sequence the human genome — our essential code for life — had whetted public appetite for medical miracles in the diagnosis, treatment and prevention of even the most complex of common diseases. "Imagine a world where families leave the hospital with their newborns and take their baby's complete genetic profile with them on a CD-ROM," Venter told his audience. "And imagine a world where your physician has as part of your medical record your genetic code, which can be used to determine, for example, your risk profile for side effects from drugs or other medical treatments. These might be possible in a genomics-based medical system in the near future." (Massoglia 2003)

"Imagine a world..." If the case of phenylketonuria (PKU) is any guide to our imagination, significant complexities should be expected to arise if neonatal genetic diagnosis and advice about risks and possible protective measures become widespread. Until the 1960s people with two PKU genes (i.e., homozygous) always suffered severe mental retardation. But now the brain damage can be averted by a special diet free of the amino acid phenylalanine following detection of those newborns having high phenylalanine levels. Yet, as Paul's (1998) history of PKU screening describes, the certainty of severe retardation has been replaced by a chronic disease with a new set of problems. Screening of newborns became routine quite rapidly during the 1960s and 70s, but there remains an ongoing struggle in the USA to secure health insurance coverage for the special diet and to enlist family and peers to support PKU

individuals staying on that diet through adolescence and into adulthood. For women who do not maintain the diet well and become pregnant, high phenylalanine levels adversely affect the development of their non-PKU fetuses; such "maternal PKU" is a public health concern that had not previously existed.

In contrast with the picture of environment overcoming genetic determination, PKU individuals are subject to heterogeneous influences on their pathways of development over the life course. A person with PKU who complies with the special diet may have experienced one or more of the following during their upbringing: health insurance coverage of the diet, a family that went without other amenities so as to afford it without insurance, or parents affluent enough to afford it anyway; a family that bent their cultural or religious dietary norms to accommodate the diet or a family without such norms; parental support for resisting peer pressure to eat other foods and/or a school community that instilled sensitivity to special needs; a sense of responsibility as a female in relation to the adverse effects of high levels of phenylalanine if they become pregnant on the development of their fetuses; summer camps where they meet other teenagers with PKU; and so on.

Another way of looking at the more complex picture is that development involves questions about control and social infrastructure and opens up possibilities for participation. Who is responsible if a baby is diagnosed with PKU, protective measures are not taken or are not sustained, and the child becomes a retarded adult or mother of a child with maternal PKU? Anyone wanting to improve the lives of PKU individuals needs to consider where they are prepared to get involved—Would the best point of engagement be around reduction in false positives or negatives? Diagnosis of variability in effects of exposure? Personal motivation and understanding of people with some mental deficits? Support groups for individuals and families? Insurance coverage for the special diet and for counseling? Paid family leave, or…? The possibilities for participation are diverse, depending on how people who want to help… can build or adjust the relevant social infrastructure (Taylor 2009a).

In short, the common claims that molecular biology and biotechnology will allow genetic information to reshape human life are fantasies in the sense that in practice many diverse materials, tools, and other people have to be engaged to realize any enduring result (Robinson 1984).

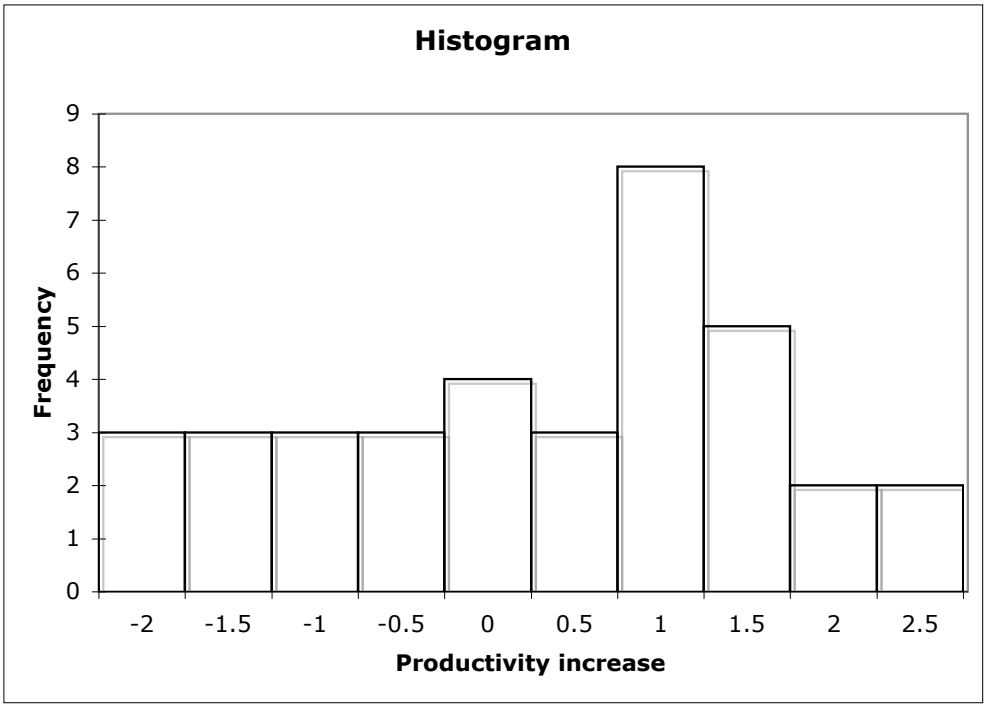
Heterogeneity #4, Deviation from the type -> #10, Participatory restructuring of the dynamics that generated the data.

(What follows is from an unfinished 2008 thought-piece)

While preparing to teach a course on epidemiology for non-specialists I made a websearch for a simple teaching example on the t-test for comparing the means (averages) of two groups for some measurement. The first example I found compared the mean productivity for two groups of workers, one group of 40 workers averaging 4.8 (in some unspecified units) with a standard deviation of 1.2 and the

other group of 45 averaging 5.2 a standard deviation of 2.4. Thinking about this example led me to articulate the sequence of thoughts and questions that follow about the foundations of statistical analysis. In particular, my inquiry explores contrasts between: the statistical emphasis on averages or types around which there is variation or noise; variation as a mixture of types; the dynamics (or heterogeneous mix of dynamics) that generated the data analyzed; and participatory restructuring of these dynamics in the future. A key issue is who is assumed to be able to take action—who are the "agents"—and who are the subjects that follow directions given by others. …[Basic sections on t-test omitted here]

3. There is something else I didn't yet mention: in the original example there was actually only one workplace—the first group in the example is made up of workers measured on one day; the second group is made up of workers measured on a later day when the music was playing. The different size of the groups is simply related to different numbers of missing measurements on the two days. We could, therefore, look at the change in productivity for individual workers who were measured on both days. Suppose that we go back to the first example and find that this change averaged 0.5 with a standard deviation of 1.3 for the 36 workers measured on both days (Figure 2). The chance of a mean difference of this size if the workers actually came from the same population—that is, if music playing had no systematic effect on individuals' productivity, whether good or bad—is 0.01… Given that the mean difference is positive, again the obvious thing to do is for the employer to play the music.



4. Yet, given that the mean difference is 0.5 and the standard deviation is 1.5, there must be many individuals who show a negative difference, that is, whose productivity declined when music was playing. In fact, this was the case for 12 of the 36 (see Figure 2). Should they oppose the playing of music, even though they are in the minority? If they do, should the employer ignore their opposition given that the firm's average individual productivity increases? Does the employer have to power to ignore any opposition? If so, the employer's power to switch on the music comes at the expense of one third of the workforce. In effect, the employer treats them as part of a music-enhances-productivity population, even though they don't fit this type.

5. The employer, faced with competition from other firms and cognizant of obligations to shareholders, might justify playing music by pointing to the increase in average productivity of the workers, which translates into an increase in overall productivity of the firm. There are, however, other paths to higher overall productivity that the employer could consider. The employer might start by asking individuals in the minority why their productivity decreased when the music played. Suppose it turned out that the tasks of those whose productivity decreased required greater concentration than the tasks of their fellow workers, or that the music chosen is not to their liking. The employer might then rearrange the workplace so that music was not played in areas where workers had to concentrate hard. Or, using headphones linked to airplane-style audio-systems, individual workers might choose from a selection of musical styles. Once the employer starts consulting individual workers, the employer might go on to ask individuals whose productivity increase was well above the mean increase to explain why. It might turn out, for example, that the music countered the tedium of their work and made them less likely to take extended bathroom breaks. By learning about the different individuals, the employer is able, in effect, to dividing the range of individuals into a set of types in relation to working when music is playing. Actions taken by the employer can then be customized accordingly. Such actions might even lead to a higher overall productivity for the firm than switching on music for all. Of course, switching on music for all is simpler and probably less expensive, but it is a matter of empirical investigation whether the firm's net profit would increase more through the customized changes or the simpler one-size-for-all action.

6. There are other things to consider about the one-size-for-all action by the employer. It keeps our focus on productivity in relation to playing music or not, and thereby keeps attention away from the dynamics (or mechanisms or causal connections) through which factors in addition to music influence productivity. We are left to hope that whatever the dynamics are, the addition of music does not lead to any long-term shifts in them. In other words, whatever dynamics generated the data we analyze, we assume that these same dynamics continue into the future even after playing music is added to them. Perhaps, however, a number of workers, including even some who like music, react negatively to the employer exerting the power to pipe in music, worrying, say, that this opens the door to advertizing, anti-union messages, and so on. Moreover, to some extent, a similar assumption about the continuation of

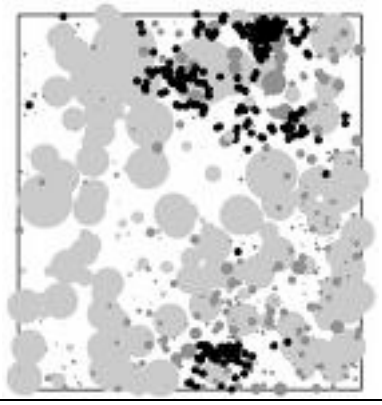
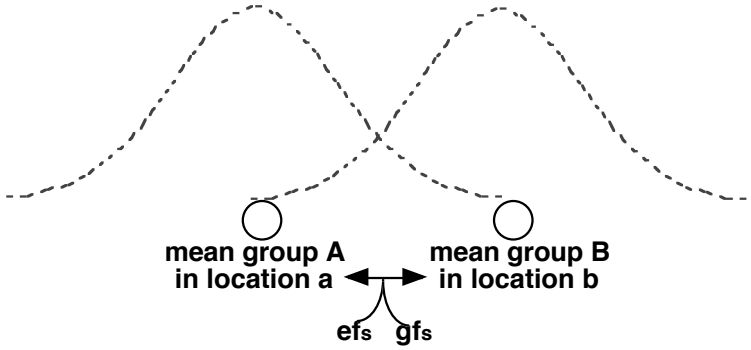
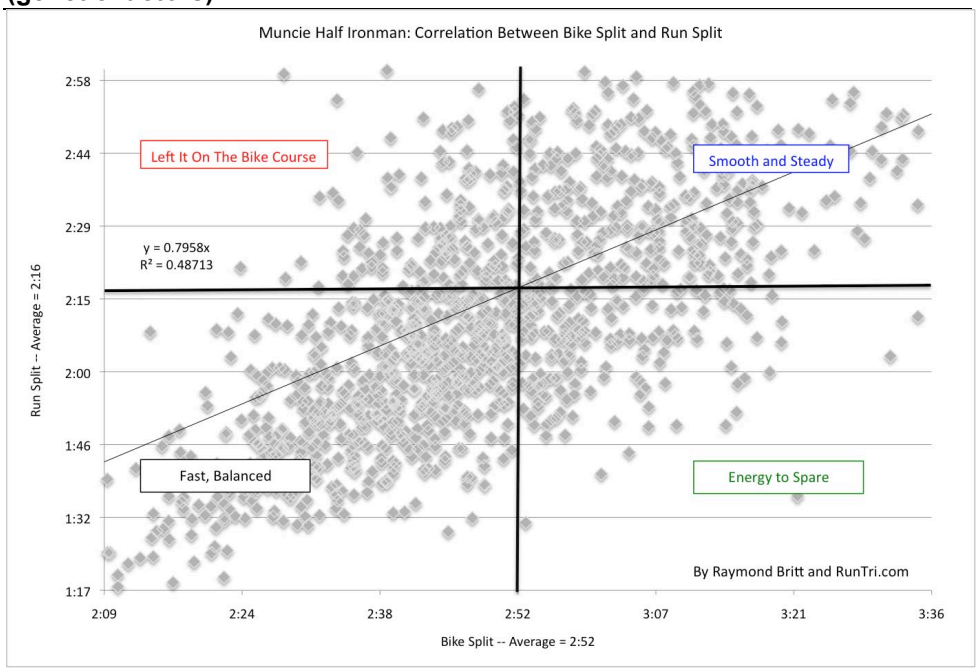
past dynamics underlies the customized actions. For example, if headphones were used so as to allow choice of music, would the quality of intra-office communication continue as before? However, there is one difference between the one-size-for-all and customized actions. The latter, by acknowledging the range of circumstances underlying the increases and decreases in individuals' productivity, opens the door to further attention to the dynamics through which factors in addition to music influence productivity. Of course, much more data is needed to investigate these dynamics and the employer might judge as unwarranted the cost of collecting and analysing the data and acting on any results.

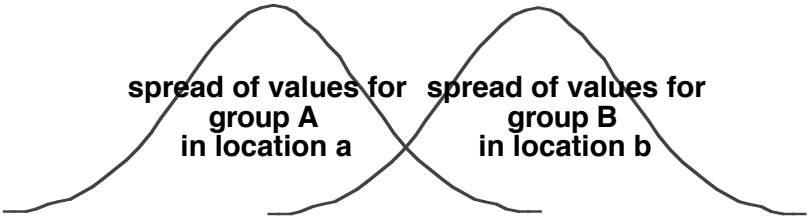
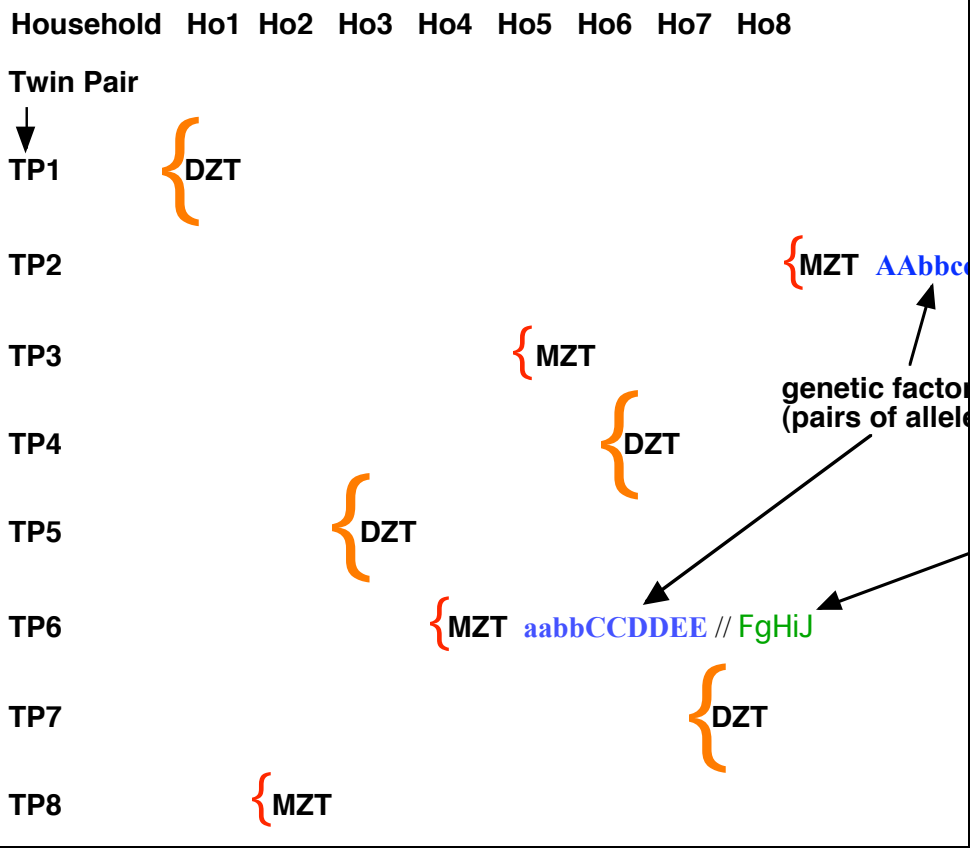
7. Imagine, however, an employer who consults workers, acknowledges the range of circumstances influencing productivity, and worries about whether past dynamics continue even after an intervention (here: switching on music) into them. These steps open the door to the employer mobilizing the workers in a participatory planning process. Skilful facilitators can lead participants through processes that elicit diverse items of knowledge about the current circumstances, generate novel proposals for improvement, and ensure that the participants are invested in collaborating to bring the resulting plans to fruition (Stanfield 2002). If this collaborative change happens, it would matter less whether the past dynamics continued as before because the workers would have become agents in the ongoing assessment and reorganization of their work lives. Moreover, improvement in productivity could result from plans unrelated to the initial issue about having music played. Of course, this scenario assumes that the employer and workers can all be brought together and kept interacting despite differences and tensions until plans are developed in which all are invested…

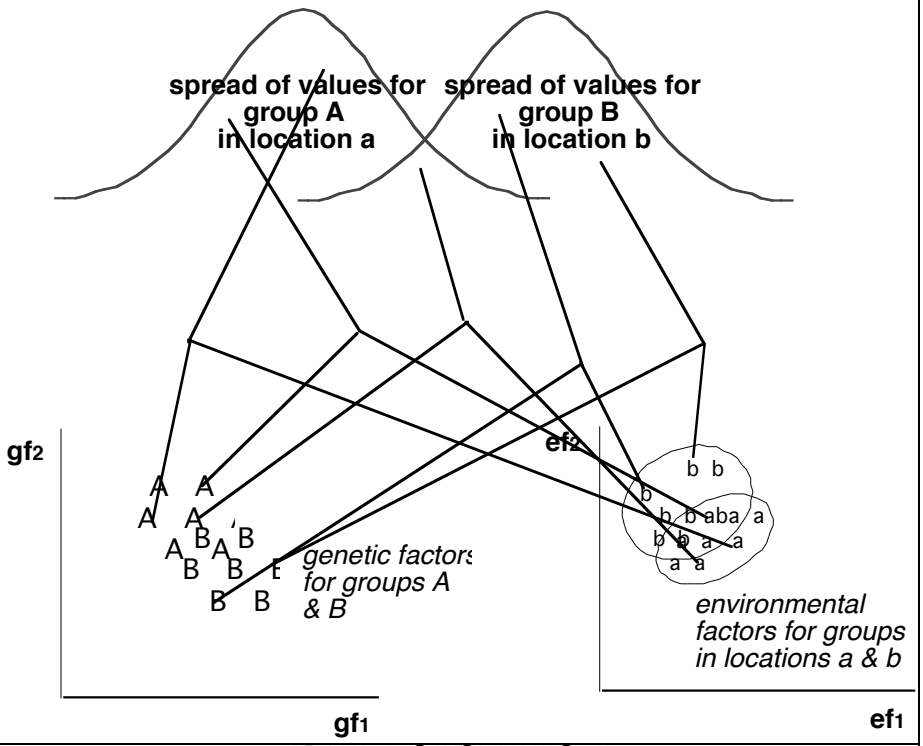
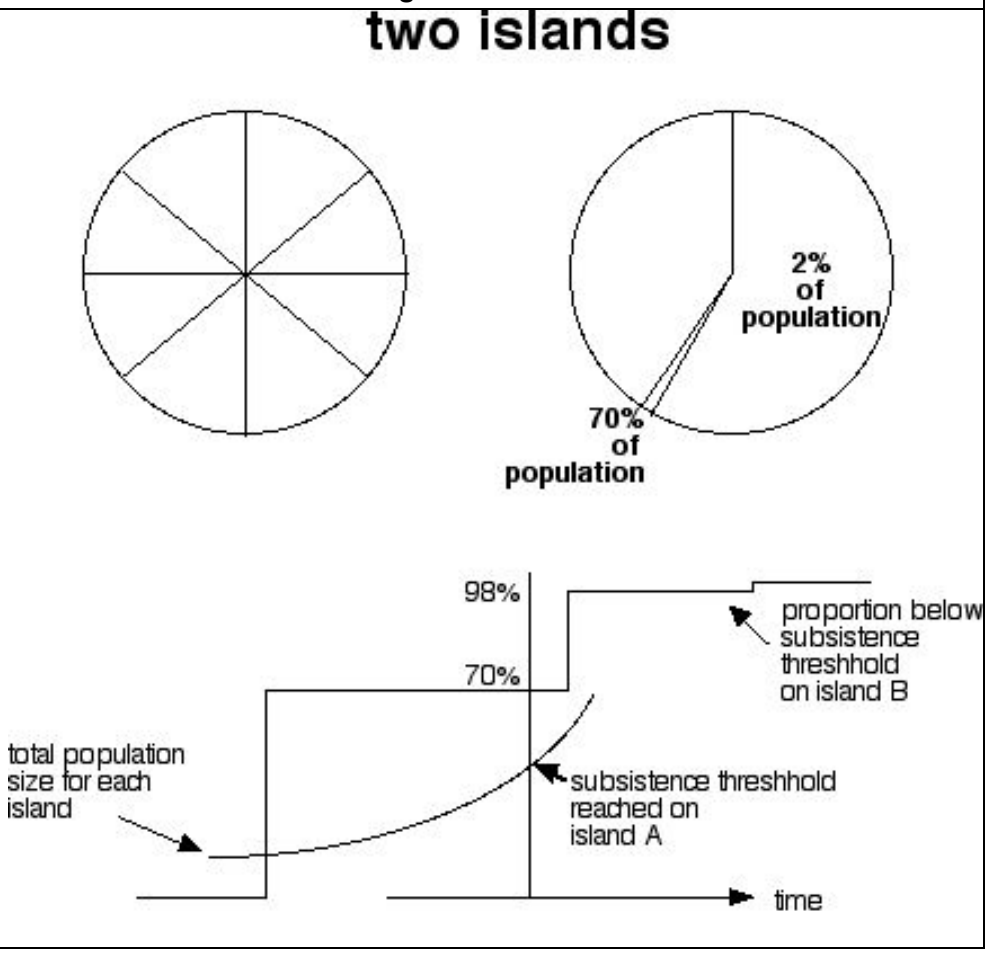
III. Images

Images—not all self-explanatory—that illustrate the taxonomy of Section 1.

STATIC 1. There is an assortment, each a separate type ("cabinet of curiosities")	

2. Mixture of types (e.g., allelic heterogeneity & locus heterogeneity in genetics)	
3. Trait = composite of types	(analogy: the 3 components of a triathalon)
VARIATIONAL 4. There is noise or error, but that is deviation from the type or essential trajectory	 **difference in means explained by efs (environmental factors) and gfs (genetic factors)**  **source: http://www.runtri.com/2011/07/muncie-half-ironman-703-results.html**

5. Variation in a set of traits involves a composite of variance/covariance structures (statistical heterogeneity)	
6. There is variation, not types	
7. When similar responses of different individual (e.g., genetic) types are observed, it is not necessarily the case that similar conjunctions of risk or protective factors have been involved in producing those responses (=possibility of "underlying heterogeneity")	

	
DYNAMIC 8. Variation produces qualitative changes in results from standard theory based on uniform units (e.g., theory about Malthusian population growth, tragedy of the commons, prisoner's dilemma)	

9. Heterogeneity in pathways of development —Variants from Taylor (2005):

9a. "Intersecting processes" Processes operating at different spatial and temporal scales that cut across the boundaries of the situation under consideration and restructure its "internal" dynamics.

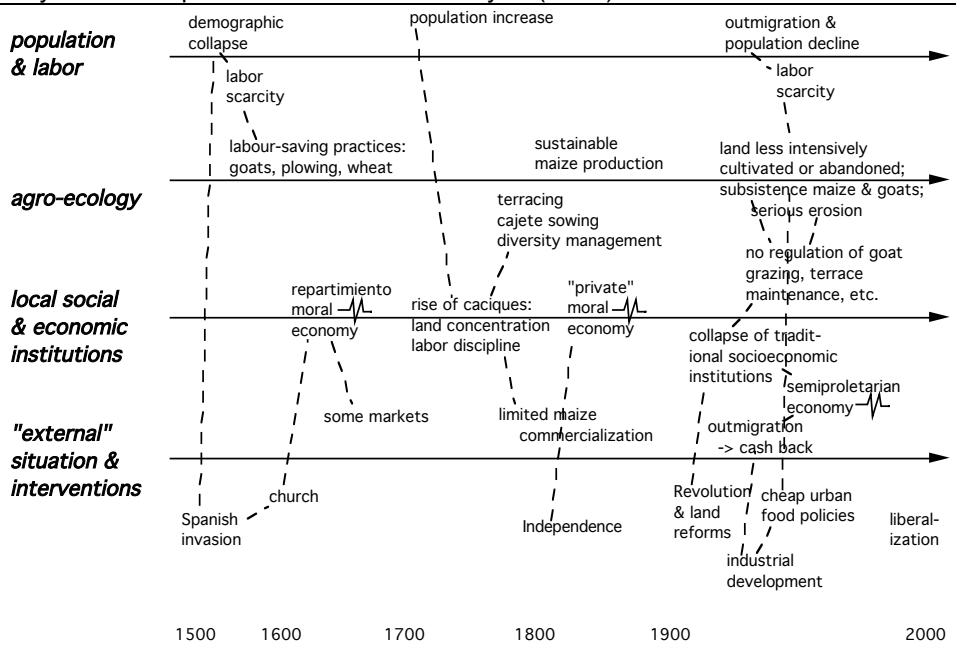


Figure 5.6 (from Taylor 2005). Intersecting processes leading to soil erosion in San Andrés, Oaxaca… The dashed lines indicate connections across the different strands of the schema. The zig-zag lines indicate institutions that rely on relationships of inequality.

9b. "Unruly complexity," which arises whenever there is ongoing change in the structure of situations that have built up over time from heterogeneous components and are embedded or situated within wider dynamics.

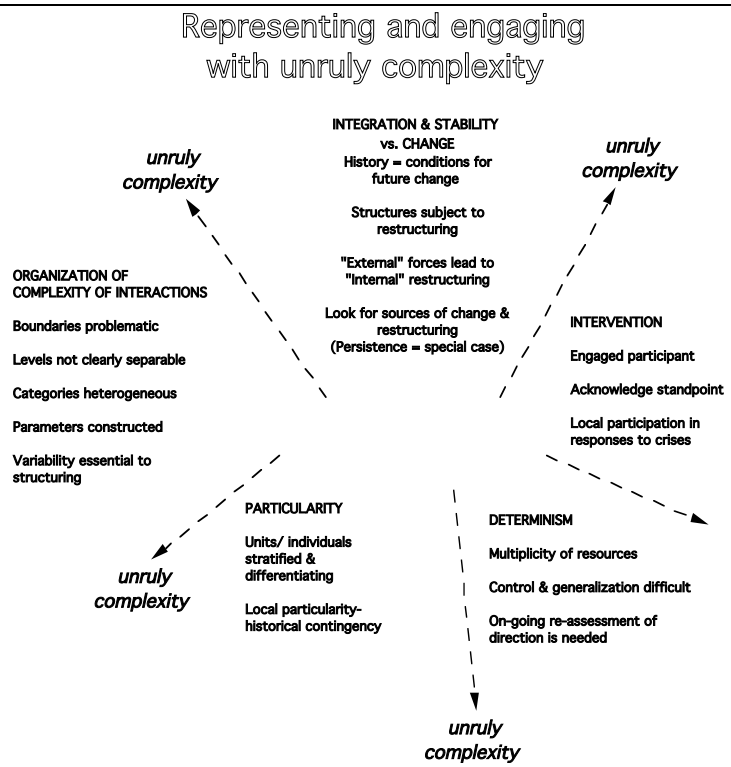
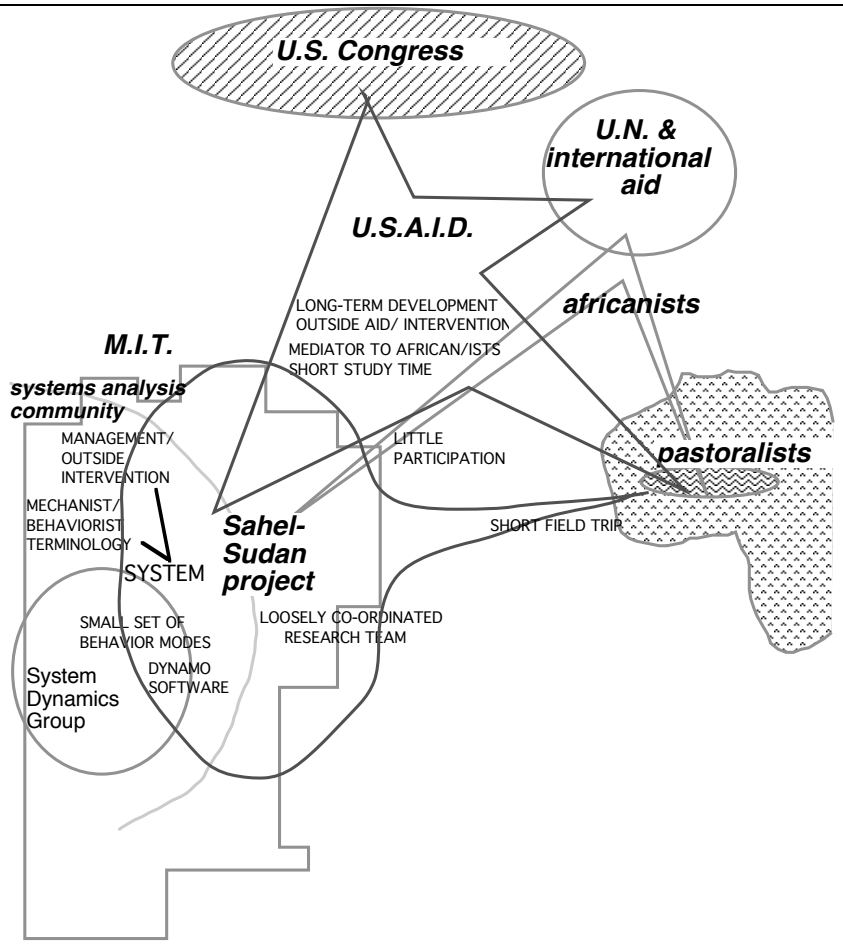
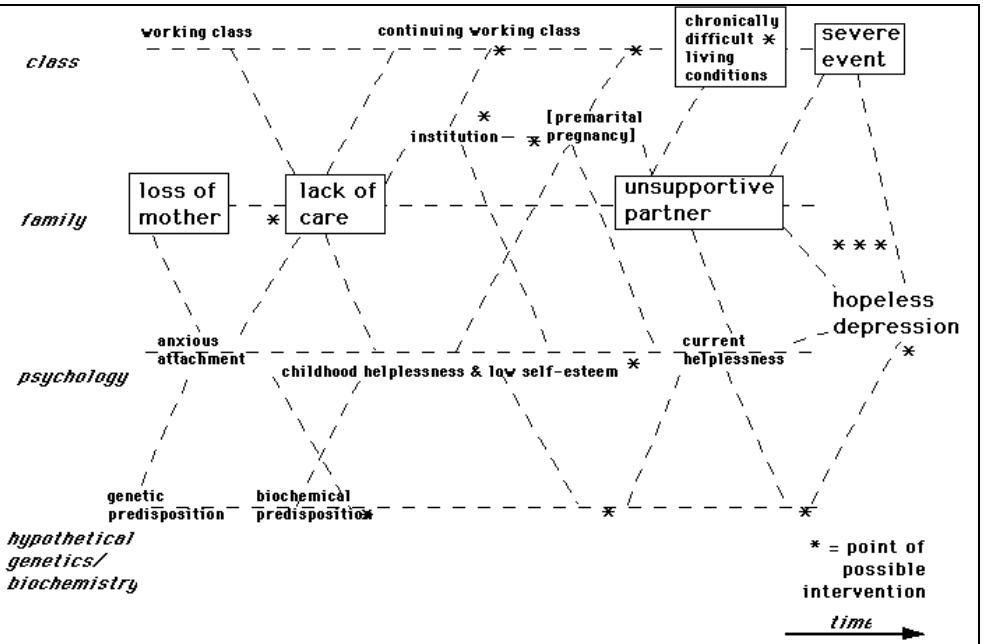
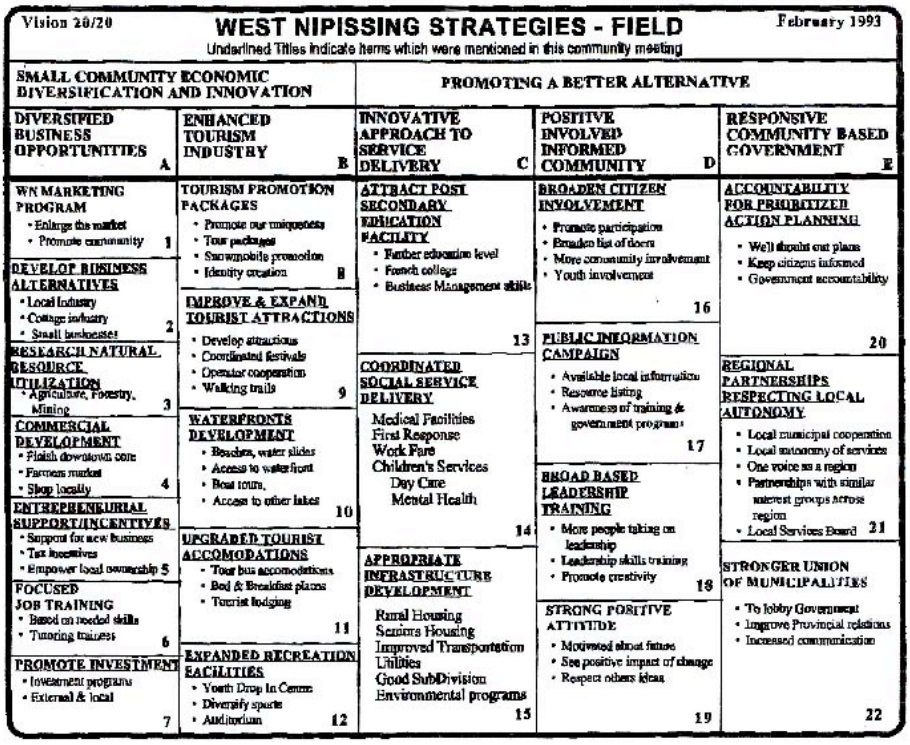


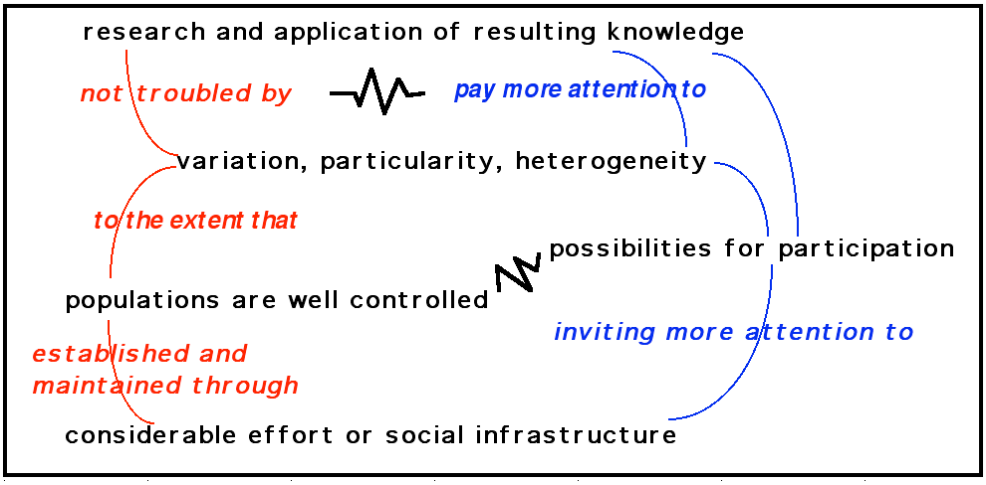
Figure 5.5 (from Taylor 2005) Conceptual and practical moves that help researchers represent and engage with the unruliness of complexity

9c. "Heterogeneous construction," in which, researchers establish knowledge and technological reliability through practices that are developed through diverse and often modest practical choices. This is the same as saying the researchers are involved in contingent and on-going mobilizing of diverse materials, tools, people, and other resources into webs of interconnected resources.	 Figure 4.6 (from Taylor 2005) An impressionistic schema depicting diverse agents and selected resources involved in the construction of Picardi's system dynamics models. The size of the elements signifies their relative importance.

DYNAMIC-PARTICIPATORY 10. Participatory restructuring of the dynamics (intersecting processes, unruly complexity, or heterogeneous construction) that generated the data.	 Pathways to severe depression in a study of working class women (discussed in Taylor 2009a, building on Brown and Harris 1989). The dashed lines indicate that each strand tends to build on what has happened earlier in the different strands. * indicate possible points of intervention that would modify the intersecting processes.
11. Participatory restructuring through multiple points of engagement, which occurs in tension with deployment or withholding of trans-local knowledge and resources.	

IV. Coda: Heterogeneity and Control

Several of the vignettes speak to a broad contention I would make about heterogeneity and control: In relation to modern understandings of heredity and development over the life course, research and application of resulting knowledge are untroubled by heterogeneity to the extent that populations are well controlled. Such *control* can be established and maintained, however, only with considerable effort or *social infrastructure*, which invites more attention to possibilities for *participation* instead of control of human subjects. On the control side, people can be made to fit types in many ways: through stereotyping, screening and surveillance, population health measures, diagnostic manuals in psychology, reassignment surgery, ignoring non-conformers, and so on. On the participation side, Taylor (2005) describes diagramming of intersecting processes to expose multiple points of engagement, "mapping" by researchers of the complex situations they study and their own complex situatedness, and well-facilitated participatory processes.



Does the contention about heterogeneity and control make sense in data analysis? Does it have relevance beyond heredity and life course development?

V. References

Borges, J. L. (1964). Other inquisitions, 1937-1952. Austin: University of Texas Press.

Brown, G. W. and T. O. Harris, Eds. (1989). Life Events and Illness. New York: Guilford Press.

Buja, A. (1996). "Interactive graphical methods in the analysis of customer panel data: Comment." Journal of Business & Economic Statistics 14(1): 128-129.

Colquhoun, J. (1998). "Why I Changed My Mind About Water Fluoridation." Fluoride 31(2): 103-118.

Cook, D. and D. F. Swayne (2007). Interactive and Dynamic Graphics for Data Analysis. New York: Springer.

Downes, S. M. (2004). Heredity and Heritability. The Stanford Encyclopedia of Philosophy. E. N. Zalta.

Eikelboom, J. W. and G. J. Hankey (2003). "Aspirin resistance: a new independent predictor of vascular events?" Journal of the American College of Cardiology 41: 966-968.

Foucault, M. (1971). The order of things: an archaeology of the human sciences. New York, Pantheon.

Hirschman, C., R. Alba, R. Farley (2000). "The Meaning and Measurement of Race in the U.S. Census: Glimpses into the Future." Demography 37(3): 381-393.

Kaplan, J. (2000) The Limits and Lies of Human Genetic Research. New York: Routledge

Massoglia, M. P. (2003). "Genomics and 'The Promise of Tomorrow'." Visions (Wake Forest University School of Medicine) Winter/Spring.

McLaughlin, P. (1989). "Obstacles to a new sociology of agriculture: The persistence of essentialism." Working Paper, Department of Rural Sociology, Cornell University.

Nuffield Council on Bioethics (2002). "Genetics and Human Behavior: The Ethical Context." http://www.nuffieldbioethics.org viewed 22 June 2007.

Paul, D. (1998). The history of newborn phenylketonuria screening in the U.S. Promoting Safe and Effective Genetic Testing in the United States. N. A. Holtzman and M. S. Watson. Baltimore: Johns Hopkins University Press: 137-160.

Robinson, S. (1984). "The art of the possible." Radical Science Journal 15: 122-148.

Rose, G. (2008 [1992]). Rose's Strategy of Preventive Medicine. Oxford: Oxford University Press.

Rutter, M. (2002). "Nature, nurture, and development: From evangelism through science toward policy and practice." Child Development 73(1): 1-21.

Stanfield, R. B. (2002). The Workshop Book: From Individual Creativity to Group Action. Toronto: Canadian Institute of Cultural Affairs.

Taylor, P. J. (2005). Unruly Complexity: Ecology, Interpretation, Engagement. Chicago: University of Chicago Press.

Taylor, P. J. (2008). "Why was Galton so concerned about "regression to the mean"?—A contribution to interpreting and changing science and society." DataCrítica 2(2): 3-22.

Taylor, P. J. (2009a). "Infrastructure and Scaffolding: Interpretation and Change of Research Involving Human Genetic Information." Science as Culture 18(4): 435-459.

Taylor, P. J. (2009b). "Nothing Reliable about Genes or Environment: New Perspectives on Analysis of

Similarity Among Relatives in Light of the Possibility of Underlying Heterogeneity." Studies in History and Philosophy of Biological and Biomedical Sciences 40(3): 210-220.

Taylor, P. J. (2010a). "Three puzzles and eight gaps: What heritability studies and critical commentaries have not paid enough attention to." Biology & Philosophy 25(1): 1-31.

Tilley, N. 2000. “Realistic Evaluation: An Overview” (Presented at the Founding Conference of the Danish Evaluation Society, September 2000), http://www.evidence-basedmanagement.com/research_practice/articles/nick_tilley.pdf, viewed 29 August 2011.

Weldon, K. L. (2000), "A Simplified Introduction to Correlation and Regression," Journal of Statistics Education, 8, http://www.amstat.org/publications/jse/secure/v8n3/weldon.cfm, viewed 22 Jun '09.

Wright, S. (1920) cited in Provine, W. (1986). Sewall Wright and Evolutionary Biology, Chicago: University of Chicago Press, p. 154.