CHAPTER 3. SCIENTIFIC EXPLANATION IN THE LARGE

Once a science has organized its basic predicates and the primary *principles conjectured to subsume them (1st-grade systemacy) into conceptions of variables governed by functional *laws (2nd-grade systemacy), it is ready to undertake more advanced (3rd grade) integrations of scientific knowledge, both horizontally and vertically, in which our explananda are no longer just single events but states or trajectories of complex systems and laws themselves. Horizontal systematizations comprehend expanses of causal networks--events richly interconnected by production transitivities--by compounding causal recursions into accounts of process regulation, growth and, at the apex of scientific insight into how things work, system dynamics. And vertical systematizations explain how macro-phenomena derive acausally from collections of micro-events.

Be forewarned, however, that the formalisms that achieve these complex integrations demand careful study: You cannot rush through their ensuing exposition and expect your commonsense intuitions to cream off its gist. To be sure, not everything below calls for your immediate attention; in fact, several short passages and one long subsection are marked with double brackets (⟦/⟧) to indicate that these are technicalities best omitted on first reading. But formalistic esoterica are the very heart of what is at issue here; and to appreciate what is being said, you must think through the notation and savor the logical force of what these rarified abstractions would assert were their schematic terms fleshed out by full-blooded conceptions of specific domains/variables/abstractors/translocators/transducers.

Considering this Chapter's relentless stacking of one formalistic construction upon another with scarcely any illustrations to root them in your own experience, I had best preview the ground to be covered and tell what may be in this for you. The material builds in five waves (with a nontechnical epilog), starting with the lawfulness of causal processes. This beginning (pp. 67-80) is fairly standard stuff in that it merely exhibits the SLese essence of a specialized, enormously powerful way to explain/predict the behavior of causal systems that has long enjoyed extensive applications in the more advanced quantitative sciences. It is important for you to comprehend the distinctive format of process laws if you want to understand the hard-science conception of how nature works, and even more so if you aspire to excellence as a practicioner of science. (Professional science training by no means assures this understanding; in particular, education in psychology profoundly neglects causal recursion in its near-exclusive emphasis on Analysis-of-Variance research designs, with the stultifying result that few modern psychologists ever learn to think system-dynamically.) Even so, this section's technicalities are not explicitly foundational for what comes after; so without penalty you can skim or omit whatever details here exceed your tedium tolerance.

Secondly (pp. 81-88) comes an easy overview of assorted causal metaprinciples in accord with which some well-SLesed laws (or *laws) follow from others. This brief section is a major contribution to the theory of lawfulness; for with the notable exception of Mediated Composition (the main metaprinciple of causal recursion), scarcely any of these derivations have heretofore been recognized for what they are, despite the ubiquity of their use in technical practice. The importance of making these explicit is twofold: For one, as demonstrated by numerous subsequent citations, their recognition is required to make clear the logic by which ensembles of micro-laws generate lawful molar systems. And secondly, verbalizing them forces awareness upon us of seminal problems in the nature of molar causality that cry for deeper inquiry. We cannot pursue those particular issues here; but even to acknowledge their challenge is a significant step ahead.

With Metaprinciple preparations in hand, we next turn to the SLese conception of how a macro-system's molar behavior supervenes upon the assembly structure and micro-causal functioning of its parts. This is developed in three phases which, regrettably but inescapably, are all formalistically formidable. First (pp. 88-98) comes an account of how modern methods of data analysis fit statistical models to sample observations. The schematisms in which I couch this overview will seem strange to readers familiar with the multivariate literature; for I emphasize SLese operations that are largely concealed by the conventional algebra whose elliptic notation for these models is designed for efficiency in mathematical analysis. However, our concern here is not to enhance extant multivariate methodology (though exhibiting its deeper SLese character has considerable potential for that), but to observe how these well-established technical practices are a major paradigm of molar explanation. The acausal supervenience of a statistical sample's holistic features/behavior upon that of its individual members is utterly transparent; and once it becomes clear (if one can surmount the formalisms) how the repertoire of SLese constructions comes together in impressively nontrivial real-life reductive explanations of statistical wholes by their parts--honestly verbalizable statements with determinate substantive content, not just promissory hand-waving--we are set to appreciate how this very same SLese logic of supervenience applies to macro-objects having compositions vastly more intricate than that of statistical aggregates.

In this next phase (pp. 98-110) of development--the molar spin-off of complex micro-structure--we break new ground: We are no longer just clarifying the logic of micro-molar stories already common in the extant scientific literature, but are instead outlining a theory of how such accounts could be given for causally cohesive macro-objects were our knowledge of their micro-constituents sufficiently complete. It is far from certain that this theory's present statement expressly schematizes all salient micro-determinants of molar phenomena; and indeed I will acknowledge important respects in which it is rudimentary. You will find little in this that you can use for substantive applications, and that may be reason enough for you to give it short shrift. But if you are serious about the nature of supervenience, levels of organization, molar causality, reduction/emergence, and their like, you should feel some obligation to struggle with Def. 2 and its molar consequences. For to my knowledge this is the first conceptually articulate model to appear on how macro-objects with strong internal structure derive their holistic lawfulness from their micro-causal underlay. You may be unimpressed with this account's depth; but it sets benchmarks for you to criticize or improve upon if you care about this classic cluster of obscurities.

Indeed, primitive as this pioneering probe of supervenient macro-causality may be, it suffices to expose certain deep puzzles, idealized out of existence in the extant systems-theoretic literature, of how stable molar lawfulness in real-world macro-subjects can obtain at all much less be humanly comprehensible. These are turgid technicalities (or at least I manage to make them so) that only the most dedicated specialists in System Structure will want to pursue. Nevertheless, they ground such important issues in causal ontology, in particular, the nature of structure-transcending "functional" properties and the supervenience shaping of infra-causal essences into molar-causality accidents that it seems only proper to close out our discussion of molar explanation with a synopsis of the structural-variation problem, its abstract formal solution, and what it takes for that solution—which on close inspection looks hoary—to be sometimes in fact a practical one. Pp. 110-123 attempts just that as a first sketch awaiting elaboration.

Scientific systemacy 3H. Causal recursion and system dynamics.

A "process" is essentially one damned thing after another, i.e., more soberly, a sequence of simple or compound events in which each stage is causally consequent upon its precursors. To predict/explain a process we must appeal to an ensemble of laws that subsumes each mediating event in the sequence as output under one law and input under another. Detailed analysis of process regulation and system dynamics requires careful attention to connectivity relations among the locations of participant events, as provided for under precondition placeholder $\mathcal{T}$ in law schema (9'); and what formalisms are most generically insightful for study of system behavior is still not entirely clear to me. Even so, it is decently straightforward to formalize causal recursions among laws that have been put into form (8) by a suitable embedding of their locus structures in the t-derivational character of their manifest variables. Imposition of form (8)--which places no constraint on the complexity of translocators variously included in $\underset{\wedge}{\underline{y}}$ and the components of $\underset{\wedge}{\underline{X}}$ --appears to lose generality at most in the breadth of domains and perhaps not even severely in that.

Causal recursions, system dynamics, and process regularities are all slightly different technical perspectives on <u>causal continuance</u>: Causal effects generally produce consequences of their own, thereby creating causal progressions wherein events earlier in the sequence influence later ones through the mediation of others more proximate to them. To see how the longer-reach lawfulness in such sequences derives from that of its intervening steps, use ellipsis (8') for law-form (8) to suppose that

$$\underline{L}_1\text{: In } \underline{D}_1,\ \underline{z}' = \phi_1(\underline{X})\ , \qquad \underline{L}_2\text{: In } \underline{D}_2,\ \underline{y} = \phi_2(\underline{z}',\underline{Z})\ ,$$

are two laws whose domains have a non-null intersection $\underline{D}_{12} = \underline{D}_1 \cap \underline{D}_2$. Then for any $\underline{o}$ in $\underline{D}_{12}$ it follows that

$$\underset{\wedge}{y}(\underline{o}) = \phi_2(\underset{\wedge}{z}'(\underline{o}),\underset{\wedge}{Z}(\underline{o})) = \phi_2(\phi_1(\underset{\wedge}{X}(\underline{o})),\underset{\wedge}{Z}(\underline{o}))\ .$$

That is, since the output $\underset{\wedge}{z}'$ of $\underline{L}_1$ is a component of $\underline{L}_2$'s input $[\underset{\wedge}{z}',\underset{\wedge}{Z}]$, $\underline{L}_1$ and $\underline{L}_2$ integrate into an overarching law $\underline{L}_{12}$, namely,

$\underline{L}_{12}$: In $\underline{D}_{12}$, $\underline{Y} = \phi_{12}(\underline{X},\underline{Z})$ $\quad (\phi_{12}(\_\_,\_\_) =_{\text{def}} \phi_2(\phi_1(\_\_),\_\_))$,

under which the components of compound variable $[\underset{\wedge}{X},\underset{\wedge}{Z}]$ jointly determine variable $\underset{\wedge}{Y}$ in domain $\underline{D}_{12}$ through the (partial) mediation of variable $\underset{\wedge}{z}'$. I shall call this $\underline{L}_{12}$ the *product* of integrable law-pair $\langle\underline{L}_1,\underline{L}_2\rangle$.

Example. Suppose that $\underline{D}$ is a population of human perceivers at times when they visually focus on a spot $\underline{S}_1$ of variable luminance surrounded by a darker field $\underline{S}_2$. Let $\underset{\wedge}{y}$ be the subjective brightness of $\underline{S}_1$ for the perceiver on some psychophysical rating scale; let $\underset{\wedge}{z}_k$ ($\underline{k} = 1,2$) measure the perceiver's neural activity at the retinal region stimulated by light from $\underline{S}_k$; and let $\underset{\wedge}{x}_k$ ($\underline{k} = 1,2$) be the photometric intensity of light leaving $\underline{S}_k$. Finally, suppose that for certain positive numerical constants $\langle\underline{a},\underline{b},\underline{c},\underline{d},\underline{r}\rangle$, retinal illumination from an external stimulus determines retinal activity according to

$\underline{L}_{ak}$: In $\underline{D}$, $\underline{z}_k = \underline{a}(\underline{x}_k - \underline{b})^{r}$ $\quad (\underline{k} = 1,2)$,

while

$\underline{L}_b$: In $\underline{D}$, $\underline{y} = \underline{cz}_1 - \underline{dz}_2$

is how contrastively patterned retinal excitation gives rise to perceived focal brightness. Then through the mediation of $\underset{\wedge}{z}_2$, field-luminance $\underset{\wedge}{x}_2$ conjoins retinal activity $\underset{\wedge}{z}_1$ to determine subjective brightness $\underset{\wedge}{y}$ according to the product of $\langle\underline{L}_{a2},\underline{L}_b\rangle$, namely,

L_c: In D, $y = cz_1 - da(x_2 - b)^r$.

And from integration of $\langle L_{a1}, L_c\rangle$, in turn, distal stimulus intensities $\underset{\wedge}{x}_1$ and $\underset{\wedge}{x}_2$ jointly determine perceived focal brightness $\underset{\wedge}{y}$ through the mediation of proximal excitation $[\underset{\wedge}{z}_1, \underset{\wedge}{z}_2]$ by transduction

L_d: In D, $y = ca(x_1 - b)^r - da(x_2 - b)^r$.

That is, with the detailed transducers in L_{ak} and L_b abbreviated as $\psi(\underline{\ \ }) = a(\underline{\ \ } - b)^r$ and $\phi(\underline{\ \ },\underline{\ \ }) = c\cdot\underline{\ \ } - d\cdot\underline{\ \ }$, respectively, composition of $z_2 = \psi(x_2)$ into $y = \phi(z_1, z_2)$ yields $y = \phi(z_1, \psi(x_2))$ while composition of $z_1 = \psi(x_1)$ into the latter then gives $y = \phi(\psi(x_1), \psi(x_2))$.

It is of some importance to appreciate that determination L_c of $\underset{\wedge}{y}$ by $\langle \underset{\wedge}{z}_1, \underset{\wedge}{x}_2\rangle$ in this example is a perfectly good causal regularity given that L_{a2} and L_b are, even though psychologists do not ordinarily mix distal and proximal stimuli when speaking of joint input. In present case L_c, distal stimulus variable $\underset{\wedge}{x}_2$ has nonzero weight for $\underset{\wedge}{y}$ when conjoined with proximal stimulus variable $\underset{\wedge}{z}_1$ in transduction $y = \phi(z_1, \psi(x_2))$ because $\underset{\wedge}{x}_2$ has effects on $\underset{\wedge}{y}$ unmediated by $\underset{\wedge}{z}_1$. But in the causal equation telling how all of distal and proximal variables $\langle \underset{\wedge}{x}_1, \underset{\wedge}{x}_2, \underset{\wedge}{z}_1, \underset{\wedge}{z}_2\rangle$ work conjointly to determine $\underset{\wedge}{y}$, namely,

L_e: In D, $y = \phi(z_1, z_2) + 0\cdot x_1 + 0\cdot x_2$,

$\underset{\wedge}{x}_1$ and $\underset{\wedge}{x}_2$ have zero weights because their effects on $\underset{\wedge}{y}$ in D are wholly mediated by $[\underset{\wedge}{z}_1, \underset{\wedge}{z}_2]$.

An annoying complication for the causal status of the product law L_{12} entailed by an integrable law-pair $\langle L_1, L_2\rangle$ is that the composition $\phi_{12}(\underline{\ \ },\underline{\ \ }) = \phi_2(\phi_1(\underline{\ \ }),\underline{\ \ })$ of causal transducers $\langle \phi_1, \phi_2\rangle$ in $\langle L_1, L_2\rangle$ may be not the causal transducer of $\underset{\wedge}{y}$'s determination by $[\underset{\wedge}{X},\underset{\wedge}{Z}]$ in $\underline{D}_{12}$ but only an acausal regressor (cf. p. 37 above) coincident with the other over the subrange of $[\underset{\wedge}{X},\underset{\wedge}{Z}]$-values occurrent in $\underline{D}_{12}$.[8] But so long as the causal-mediation ordering of component variables in $[\underset{\wedge}{X},\underset{\wedge}{z}',\underset{\wedge}{Z},\underset{\wedge}{y}]$ complies with standard preconditions, in particular when no component of $[\underset{\wedge}{X},\underset{\wedge}{Z}]$ is errorlessly determined in $\underline{D}_{12}$ by the others (which is a sufficient but not necessary condition), the causal transducer under which the input of $\langle\underline{L}_1,\underline{L}_2\rangle$-product $\underline{L}_{12}$ determines its output is indeed ϕ_{12}, just as intuition desires. The generic theory of conditions under which product-laws inherit the causal status of the laws they integrate is one of the most important research frontiers now open for advanced work on the logic of causality, secondary only to the obscure but vital linkage of plausible inference to causal-order presumptions and the problem of molar causality. But since this formally complex issue is easily detached from our main concerns here, I shall presume that whenever $\underline{L}_{12}$ is stated to be the product of integrable causal laws $\langle\underline{L}_1,\underline{L}_2\rangle$, the causal ordering of variables in the latter is such that $\underline{L}_{12}$ too is a causal law whose transducer is the composition of $\underline{L}_1$'s transducer into that of $\underline{L}_2$.

Evidently, the principle of law-integration just sketched can be expanded. Call any sequence $\langle\underline{L}_1,\underline{L}_2,\ldots,\underline{L}_m\rangle$ ($\underline{m}\geq 2$) of laws a <u>c(omposable)-series</u> just in case, for each $\underline{j} = 1,\ldots,\underline{m}-1$, $\langle\underline{L}_j,\underline{L}_{j+1}\rangle$ is integrable. Then by iteration of pairwise law-production, any c-series $\langle\underline{L}_1,\ldots,\underline{L}_m\rangle$ integrates into a product law $\underline{L}^*_{1m}$ whose domain $\underline{D}^*_{1m}$ is the intersection of the domains of $\underline{L}_1,\ldots,\underline{L}_m$; and any consecutive subsequence of this c-series is also a c-series that likewise

---

[8] I have explored this situation rather thoroughly in unpublished work under the working title "Complexities of mediation structure." Believe me; details you don't want.

integrates into a product law. (Constraints are needed on the causal ordering of the variables in a c-series to insure that all its product laws derive their causal transducers by composition from those of the laws they overarch; but I have already waved off this problem by allusion to "standard preconditions.") For any c-series $\langle L_1,\ldots,L_m\rangle$, let us say that the ensemble $\mathbf{L}$ of laws comprising all L_j therein together with the products of this sequence and all its consecutive subsequences is a (*simple*) *causal recursion* over domain D^*_{1m}.[9] More generally, we may say that an ensemble $\mathbf{L}$ of laws is a (*recursive*) *causal system* over domain D^* iff (*a*) $\mathbf{L}$ is closed under the law-product operation (i.e., every product of integrable laws in $\mathbf{L}$ is also in $\mathbf{L}$), (*b*) each law in $\mathbf{L}$ is either integrable with some other law in $\mathbf{L}$ or is the product of integrable laws in $\mathbf{L}$, and (*c*) D^* is the intersection of the domains of all laws in $\mathbf{L}$.

⟦If $\mathbf{L}$ is a recursive causal system, any variable that is, or is a component of, the input (output) of some law in $\mathbf{L}$ is a *local input* (*local output*) of $\mathbf{L}$; any variable that is a local input but not a local output (local output but not a local input) of $\mathbf{L}$ is a *global input* (*global output*) of $\mathbf{L}$; and any variable that is both a local input and a local output of $\mathbf{L}$ is a *mediation* or *system variable* of $\mathbf{L}$. For any global input variable x_1 of causal system $\mathbf{L}$, and any object o in $\mathbf{L}$'s domain, $\mathbf{L}$ implies the existence of, and accounts for, a process comprising $\ulcorner \underset{\wedge}{x}_1;o\urcorner$ followed by $\ulcorner \underset{\wedge}{x}_2;o\urcorner$ followed by ... followed by $\ulcorner \underset{\wedge}{x}_{r-1};o\urcorner$ followed by $\ulcorner \underset{\wedge}{x}_r;o\urcorner$, where $\underset{\wedge}{x}_r$ is a global output variable of L. In this sequence, $\underset{\wedge}{x}_2,\ldots,\underset{\wedge}{x}_{r-1}$ are mediation variables of $\mathbf{L}$; each event $\ulcorner \underset{\wedge}{x}_k;o\urcorner$ ($k=2,\ldots,r$) is caused by $\ulcorner \underset{\wedge}{x}_{k-1};o\urcorner$ conjointly with other events recognized by $\mathbf{L}$; and for each $\langle h,k\rangle$ such that $1\le h<k\le r$, $\mathbf{L}$ contains one or more laws L_j under which $\ulcorner \underset{\wedge}{x}_h;o\urcorner$ and certain other L_j-identified events conjointly cause $\ulcorner \underset{\wedge}{x}_k;o\urcorner$ through the mediation of o's standing on variables $\underset{\wedge}{x}_{h+1},\ldots,\underset{\wedge}{x}_{k-1}$.⟧

---

[9] The phrase 'causal recursion' or 'recursive causal model' is often used more loosely than this; but in all modern multivariate applications it expresses concern for ensembles of laws that are integrable in the fashion here described.

For simplicity, I have described recursive causal systems as certain collections of form-(8) laws. But to acknowledge the epistemic imperfection of our conceptions of these, we could speak instead of conjectured causal *systems whose elements are causal *laws. Even so, the dubiety of *laws is an extra burden that we can well do without when dealing with matters of this complexity. In particular, presumption of truth allows us to ignore technicalities of coherence in a conjectured *law ensemble that would otherwise demand attention.

With a bit of care for causal-order preconditions, we can allow the outputs of form-(8) *laws to be compound variables rather than singletons. That is, (8') can be expanded into

$$(8'a) \qquad \text{In } D,\ Y = \Phi(X) \qquad (\ \underset{\wedge}{Y} = [\underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_n]\ ,\ \ \Phi = \langle\phi_1,\ldots,\phi_n\rangle\ )\ ,$$

where $\underset{\wedge}{Y}$ is a tuple of basic or derivative variables whose domains include D, $\Phi(\underset{\wedge}{X})$ is compact notation for $\langle\phi_1(\underset{\wedge}{X}),\ldots,\phi_n(\underset{\wedge}{X})\rangle$, and (8'*a*) asserts that in D, $y_j = \phi_j(X)$ for each $j = 1,\ldots,n$. We note that the *j*th component ϕ_j of compound transducer Φ may assign null weight to some of the component variables in $\underset{\wedge}{X}$, namely, to ones that are not in fact sources of $\underset{\wedge}{y}_j$ unmediated in D by the other $\underset{\wedge}{X}$-components (cf. L_e, p. 68 above); but beyond that we continue to disregard niceties about the precise causal status of the transducer in (8'*a*). In any case, it is evident that some compound-output laws are integrable into product laws in the same fashion as are suitably matched single-output laws; so we permit the individual laws constituting a recursive causal system $\underset{m}{L}$ to have form (8'*a*) as well as form (8').

An outstandingly important special case of causal recursion comprises the lagged self-products of any "auto-regressive" *law whose output variable differs from one or more components of its input only by an embedded translocator. Specifically, suppose that output variable $\underset{\wedge}{y}$ in (8/8') is a locus displacement $\underset{\wedge}{y} = [\underset{\wedge}{x}_k f]$ of the *k*th component of input compound $\underset{\wedge}{X} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m]$. Making this translocation explicit while keeping notation simple by choosing $\underline{k} = 1$ and rewriting $\underset{\wedge}{X}$ as $[\underset{\wedge}{y}, \underset{\wedge}{Z}]$ (so that $\underset{\wedge}{x}_k$ and $\underset{\wedge}{x}_k\underline{f}$ become $\underset{\wedge}{y}$ and $\underset{\wedge}{y}\underline{f}$) converts (8) to

(11) For all $\underline{o}$ in $\underline{D}$, $\underset{\wedge}{y}(\underline{f}(\underline{o})) = \phi(\underset{\wedge}{y}(\underline{o}), \underset{\wedge}{Z}(\underline{o}))$,

or more compactly

(11')[9a] In $\underline{D}$, $\underline{yf} = \phi(\underline{y}, \underline{Z})$.

[9a] Here and henceforth I shall write '$\underline{yf}$', and similarly for other notationally explicit t-derivative variables, for the value on $[\underset{\wedge}{y}\underline{f}]$ for an arbitrary member of $\underline{D}$ in counterpart to the output side of schema (8'). This notation should slide easily across your eye without provoking you to notice that composition of translocation function $\underline{f}$ into an arbitrary value $\underline{y}$ of variable $\underset{\wedge}{y}$, rather than into $\underset{\wedge}{y}$ itself qua function, makes no literal sense.

This says that each member *o* of *D* has a unique *f*-successor *f*(*o*) whose value on variable $\underset{\wedge}{y}$ is determined in part by *o*'s own value on $\underset{\wedge}{y}$. I call translocator function *f* in (11) the (manifest) *excursor* of *law (11). Its role is to identify *where* a particular cause subsumed by (11) exerts its effect. That is, if *o* is the manifest locus of compound event $\ulcorner[\underset{\wedge}{y},\underset{\wedge}{Z}];o\urcorner$, (11) describes how this determines the $\underset{\wedge}{y}$-event whose manifest locus is specifically *f*(*o*) rather than any of the other individuals in *D* which also have $\underset{\wedge}{y}$-values albeit ones due under (11) to antecedent $\underset{\wedge}{y}$-events other than $\ulcorner\underset{\wedge}{y};o\urcorner$. (I say "manifest" locus here because if $\underset{\wedge}{y}$ is t-derivative, e.g. when $\underset{\wedge}{y}$ = *the*-*height*-*of*-__'*s*-*father*, we may wish to say that the manifest locus *o* of event $\ulcorner\underset{\wedge}{y};o\urcorner$ differs by translocation from its real locus.) *All* causal laws, not just auto-regressions, include excursive relations in their preconditions on the events they connect as cause and effect, nor are these always functions; but we can let the excursors of auto-regressions go proxy here for the more general case.

As already noted (cf. (6')), the most familiar instances of causal excursion are time-displacements between stages of the same enduring subject (continuant thing) under which the *f*-successor of *s*-at-*t* (= *o*) is *s* a little later than *t*. Then (11) specifies how the value of $\underset{\wedge}{y}$ at any stage of *s*'s development affects *s*'s $\underset{\wedge}{y}$-status shortly thereafter. Moreover, this temporal paradigm can easily be generalized: Relative to any form-(11) law with domain *D* and excursor *f*, we stipulate that an enduring *D*-subject is a subset of *D* that is connected by *f*. That is, *s* is an *f*-wise enduring *D*-subject iff *s* comprises just *D*-members any two of which are related by *f* or some power of *f*. ⟦When (11) is a macro-step product of a more continuous growth process, as described later, we take *f*'s "powers" $\{f^r\}$ to include decompositional factors of *f* as well as its integer powers clarified below.⟧ Then earlier and later stages of the same *s* are identified in terms of which *D*-members can be carried into which others by powers of *f*.

Evidently, if $\ulcorner \underset{\wedge}{y}; \odot \urcorner$ is a source of the $\underset{\wedge}{y}$-value of $\odot$'s f-successor, then the latter should similarly be a source of the $\underset{\wedge}{y}$-value of $f(\odot)$'s own f-successor, and so on for arbitrarily many steps of f-progression. But an important impediment to this iteration is that although excursor f in (11) is by stipulation a function whose domain D_f includes D, so that $f(\odot)$ uniquely exists for every $\odot$ in D, $f(\odot)$ is thereby guaranteed to have an f-successor of its own only if $f(\odot)$ too is in D--which may not be so, insomuch as enduring subjects eventually reach a terminal stage, notably death, at which they have no continuations covered by the laws that govern their pre-terminal stages. That is, for any integer $r > 1$, the r-fold composition f^r of f into itself, i.e. $f^r(\_\_) =_{def} f(f^{r-1}(\_\_))$, is generally only a partial function over D insomuch as f's domain D_f may not include all of its range nor is $f^r(\odot)$ necessarily in D even when it is in D_f.

To cope with this complication, let D/f be the set of all D-members having immediate f-successors in D, i.e.,

$$D/f \ =_{def} \ \{\odot\colon \odot \in D \text{ and } f(\odot) \in D\} \ .$$

Then for each $m = 0,1,2,\ldots$, recalling that by convention $f^1 = f$ and f^0 is the Identity function over D_f, we can pick out those members of D that have at least an m-fold sequence of f-successors in D by writing $D\overset{0}{/}f =_{def} D$ and from there

$$D\overset{m}{/}f \ =_{def} \ (D\overset{m-1}{/}f)/f \ = \ \{\odot\colon \ f^r(\odot) \in D \ \text{ for all } r = 0,1,\ldots,m\} \ .$$

Note that $D\overset{m}{/}f$ has the property that for any nonnegative integers $r < m$, if $\odot \in D\overset{m}{/}f$ then $f^r(\odot) \in D\overset{m-r}{/}f \subseteq D$. That is, $f^r(D\overset{m}{/}f) \subseteq D\overset{m-r}{/}f$.

Since any law whose domain includes D is also a law over any subset of D, it follows from (11) by restricting D to just those D-members that can be described as immediate f-successors of other D-members that

(12) For any $\odot$ in D/f, $\underset{\wedge}{y}ff(\odot) = \phi(\underset{\wedge}{y}f(\odot), \underset{\wedge}{Z}f(\odot))$,

or more briefly

(12') In D/f, $yf^2 = \phi(yf, Zf)$.

More generally, for each o in $\overset{r}{D/f}$, $f^r(o)$ is in D and hence, from (11),

$$\left\{\text{In } \overset{r}{D/f},\ \ yf^{r+1} = \phi(yf^r, Zf^r)\right\} \quad (\,r = 0,1,2,\ldots,m\,) \tag{13}$$

for any positive integer m not so large that $\overset{m}{D/f}$ is of necessity empty. For any $m \geq 2$, (13) is a c-series whose domain is $\overset{m}{D/f}$ (since $\overset{m}{D/f}$ is contained in $\overset{r}{D/f}$ for any $r \leq m$) and whose overarching product-law is

$$\text{In } \overset{m}{D/f},\ \ yf^{m+1} = \phi^*_{m+1}(y, Z, Zf, Zf^2, \ldots, Zf^m)\ , \tag{14}$$

where ϕ^*_{m+1} is defined recursively as

$$\phi^*_1(\_\_,\_\_) =_{\text{def}} \phi(\_\_,\_\_)\ , \qquad \phi^*_{r+1}(\_\_,\ldots,\_\_,\_\_) =_{\text{def}} \phi(\phi^*_r(\_\_,\ldots,\_\_),\_\_)\ .$$

Finally, combining (13) with the product of every consecutive subsequence of (13) gives

$$\underset{\sim}{L}_y\text{: In } \overset{m}{D/f},\ \left\{yf^{r+t} = \phi^*_t(yf^r, Zf^r, Zf^{r+1}, \ldots, Zf^{r+t-1})\right\} \quad (\,0 \leq r < r+t \leq m+1\,)\ . \tag{15}$$

Law-system $\underset{\sim}{L}_y$ is the causal recursion into which (11) unfolds for any $m \geq 1$, albeit (15) adds little to what is already explicit in (14). It sets out the simplest formal model of *process regulation*, describing how each later stage of a succession of $\underset{\wedge}{y}$-events is determined by any selected earlier stage together with the history of exogenous influences (Z-inputs) over intervening steps of the process. We can further view (15)'s sub-ensemble for any fixed r, say $r = 0$, as telling for each object o in D whose f-successors $\{f^t(o)\}$ remain in D over a sufficiently long progression $t = 1,2,\ldots$ how $\underset{\wedge}{yf}^t(o)$ develops from $\underset{\wedge}{y}(o)$ as a function of t under accumulating input disturbances. Specifically, let $\{f^t(o^*_s)\colon t = 0,1,2,\ldots\}$ be an f-succession of stages, all in D, of the same enduring subject s commencing with s's origin o^*_s at time-zero.[10] Then $y_0(s) =_{\text{def}}$

---

[10] We are here thinking of time (or any other excursive counterpart of time) as scaled locally for each enduring subject s by successor displacements from a fixed stage o^*_s of s selected to be s's "origin." That is, s-at-(time)-$t =_{\text{def}} f^t(o^*_s)$ for each non-negative integer t. If we wish, we can further stipulate that o^*_s is the f-wise "birth of s in the sense that o^*_s has no f-precursor in D; but that is not obligatory.

$\underset{\wedge}{y}$($\underline{s}$-at-time-0) = $\underset{\wedge}{y}(\underline{o}^*_s)$ is $\underline{s}$'s "initial endowment" on $\underset{\wedge}{y}$, and for any sequence $\langle \hat{\underline{Z}}_0, \hat{\underline{Z}}_1, \hat{\underline{Z}}_2, \ldots \rangle$ of $\underset{\wedge}{Z}$-values that we choose to view as a "standard" input history on $\underset{\wedge}{Z}$, the trajectory (i.e. time-course) of $\underset{\wedge}{y}$-development projected for $\underline{s}$ from its initial endowment is the estimate-sequence $\{\underset{\wedge}{\hat{y}}(\underline{s}\text{-at-}\underline{t}) : \underline{t} = 1,2,\ldots\}$ wherein

$$\hat{y}(\underline{s}\text{-at-}\underline{t}) =_{\text{def}} \phi^*_t(y_0(\underline{s}), \hat{\underline{Z}}_0, \hat{\underline{Z}}_1, \ldots, \hat{\underline{Z}}_{t-1}) \quad (\, \underline{t} = 1,2,\ldots\,) \;.$$

And the divergence of $\underline{s}$'s actual $\underset{\wedge}{y}$-trajectory $\{\underset{\wedge}{y}(\underline{s}\text{-at-}\underline{t}) : \underline{t} = 0,1,2,\ldots\}$ from this projection is explained in terms of $\underline{s}$'s more-or-less atypical $\underset{\wedge}{Z}$-history by noting what perturbations from $\{\phi^*_t(y_0(\underline{s}), \hat{\underline{Z}}_0, \ldots, \hat{\underline{Z}}_{t-1}) : \underline{t} = 0,1,2,\ldots\}$ result when the standard $\underset{\wedge}{Z}$-parameters therein are replaced by $\langle \underset{\wedge}{Z}(\underline{s}\text{-at-time-}0), \underset{\wedge}{Z}(\underline{s}\text{-at-time-}1), \ldots \rangle$.

Despite the wealth of system complexity that can be built up from (11), this is still just auto-regression at minimum. One obvious but important enrichment is to replace $\underset{\wedge}{y}$ in (11) by a compound variable. If $\underset{\wedge}{Y} = [\underset{\wedge}{y}_1, \underset{\wedge}{y}_2, \ldots]$ is a tuple of system or "endogenous" variables while $\underset{\wedge}{Z}$ is a tuple of whatever global input or "exogenous" variables, including stochastic residuals, are needed to account for system change beyond the system's own state (i.e. $\underset{\wedge}{Y}$-value), the basis of multidimensional auto-regression is

$$\text{(16)} \qquad \text{In } \underline{D}, \quad \underline{Yf} = \Phi(\underline{Y}, \underline{Z}) \;.$$

(By saying that variables Z are "input" in (16), we imply inter alia that for each $\underline{o}$ in $\underline{D}$, none of the effect of $\lceil \underset{\wedge}{Y} ; \underline{o} \rceil$ upon $\lceil \underset{\wedge}{Y} ; \underline{f}(\underline{o}) \rceil$ is mediated by any component of $\lceil \underset{\wedge}{Z} ; \underline{o} \rceil$. However, it is not always convenient to make this causal-order constraint a strict precondition in the abstract formalization of system processes.) Formally, there is scarcely any difference between (16) and (11'); and the equations and interpretations already given for (11)'s recursive unfolding apply equally (with $\underset{\wedge}{Y}$ and Φ for $\underset{\wedge}{y}$ and ϕ) to the recursive system entailed by (16). In fact, (16) can always be subsumed under (11') by viewing $\underset{\wedge}{y}$ in the latter as the cartesian product of the component variables in $\underset{\wedge}{Y}$--which is in effect what we do when conceiving each value $\underline{Y}$ of $\underset{\wedge}{Y}$ as a "(total) system state".[11] Do not be misled by (16)'s formal simplicity, however; for this is compact notation for

$$(16')\qquad \text{In } \underline{D},\ \begin{cases} \underline{y}_1\underline{f} = \phi_1(\underline{y}_1,\ldots,\underline{y}_m,\underline{z}_1,\ldots,\underline{z}_n) \\ \quad\vdots \qquad\qquad\qquad \vdots \\ \underline{y}_m\underline{f} = \phi_m(\underline{y}_1,\ldots,\underline{y}_m,\underline{z}_1,\ldots,\underline{z}_n) \end{cases} .$$

Even when transducers $\phi_1,\ldots,\phi_m$ are of very simple form (linearity being the mathematical ideal) with all exogenous variables $\{\underset{\wedge}{z}_k\}$ constant or random, the system-state trajectories $\{\underset{\wedge}{Y}\underline{f}^t(\underline{o})\colon \underline{t} = 0,1,2,\ldots\}$ that unfold from $\underset{\wedge}{Y}(\underline{o})$ under iteration of (16') can be quite complex, with the trajectory on each $\underset{\wedge}{Y}$-component $\underset{\wedge}{y}_j$ being in general a weighted composite of several curves, of which some are monotone trends that may or may not converge to asymptotes while others are cycles differing in periodicity and damping.

Model (16) has so vast a literature, both in theory and in applications, that it seems parochial to cite any one example. Yet it is worth recalling, with Garfinkel (1981, p. 53f.), the classical elegance of its application by macrobiology to prey/predator population dynamics. In this instance, each *D*-member is a bio-community *b* at some time *t*; *f*(*b*-at-*t*) is *b* at time $\underline{t}+\Delta$; and each $\underset{\wedge}{y}_j$(*b*-at-*t*) is the quantity of some animal or plant species $\underline{S}_j$ in *b* at *t*. The abundance of

---

[11] In the systems-theoretic literature, it is often unclear whether what is meant by "system" is (*a*) a recursive ensemble of laws (or *laws), (*b*) a member of the domain *D* of those laws, (*c*) an *f*-connected subset of *D*, (*d*) some subset of the events $\{[\underset{\wedge}{x};\underline{o}]\}$ in which $\underset{\wedge}{x}$ is a variable of the system and *o* a member of its domain, or (*e*) some amalgam of all of these. When one speaks of "system state," however, there is an implicit focus on the status of some particular *o* in *D* on the totality of just the system variables (*contra* global inputs/outputs), *i.e.* the ones that occur nonvacuously on both sides of equation (16). But our conception of *o* in such contexts usually functions as a universally quantified placeholder as in (11)--which is to say that "system state" is primarily just a particular value of $\underset{\wedge}{y}$ in (11'), or of $\underset{\wedge}{Y}$ in (16), viewed from concern for its causes and effects.

each species $\underline{S}_j$ in $\underline{b}$ at $\underline{t}+\Delta$ is an increasing function of the abundances of $\underline{S}_j$ and certain other species in $\underline{b}$ at $\underline{t}$ (the more rabbits there are the greater their total reproduction, and the thicker the browse the more they flourish), but is a decreasing function of abundance in other species that prey upon $\underline{S}_j$ (the more numerous the wolves the worse for rabbit survival; the denser the rabbits the greater the strain on the browse), and is also affected in identifiable ways by exogenous conditions between $\underline{t}$ and $\underline{t}+\Delta$ such as vagaries of weather and cropping or succorance from outside the system (cf. biblical wrath of God and manna from heaven). Natural cycles of prey/predator build-up and decline are captured nicely by this model, and long-term effects of various contemplated human interventions can also be well forecast by its iteration.

Our componentializing total system state as a tuple of coordinates (values) on arbitrarily many axes (component variables) of a multidimensional space of system-state alternatives, as made explicit in (16') in contrast to (11), is essential for practical theory development in natural science. Put abstractly, this need arises from limitations on the sorts of complexity we can handle in our conceptions of transducers. But at the operational level of scientific inquiry, it is evident that we never begin to make sense out of empirical phenomena until we decompose the total system at issue into subsystems whose respective causal functionings can be worked out in partial independence of the rest. This is one reason why the Turing Machine model of "computational processes" so popular in the recent cognitive-science literature is a grotesquely inappropriate paradigm for thinking about the behavior of natural systems.

When excursor $\underline{f}$ is a within-subject time displacement, (16) is with certain mild qualifications the basic classical model of system dynamics. One qualification is that (16) does not make explicit any global-output variables. But those can be included in (16) as it stands by allowing some components of $\underset{\wedge}{Y}$ to have null weight in transducer Φ, or added by conjoining (16) with an output subsystem, e.g. $X = \Psi(Y,E)$. A more significant qualification is that system dynamics is often expressed not by *growth* law (16) but by the corresponding difference-equation

$$(17) \qquad \text{In } D, \quad \Delta_f Y = \Phi(Y,Z) - Y \qquad (\Delta_f \underset{\wedge}{Y}(o) =_{\text{def}} \underset{\wedge}{Y}f(o) - \underset{\wedge}{Y}(o)) ,$$

which describes the *change* of system-state over one *f*-step in a subject whose stages are in *D*. Formulas (16) and (17) are so directly interconvertible that there seems little to choose between them; yet they challenge the philosophy of explanation to ajudicate whether system change as expressed by (17) is *due to* the growth expressed by (16), or whether, conversely, a system grows because it changes. (Were it not flagrantly digressive here, I would argue for the former.)

Moreover, (17) serves as reminder that this is a discrete-time counterpart of the *continuous-change* models that have long been the backbone of mathematical physics. The ontological status of instantaneous change, which underlies the issue of change-vs.-growth priority, is out of bounds here. Even so, the prospect of replacing (17) by a prima facie more fundamental differential equation points out that given any macro-step growth process of form (16), there may well exist for any arbitrarily large integer *m* a micro-step auto-regression for system variables $\underset{\wedge}{Y}$ of which (16) is the *m*-fold auto-product in the way that (14) is the *m*-fold auto-product of (11). (Such infinite subdivisibility of excursion steps is essentially what it means to postulate differentiable growth.) Whether strict continuity of process stages on a particular tuple $\underset{\wedge}{Y}$ of system variables is even logically possible depends importantly on details of how we define the objects in $\underset{\wedge}{Y}$'s domain. For example, were $\underset{\wedge}{Y}$ to comprise children's scores on a battery of tests that take appreciable time to administer, the developmental stage of any child *s* to which we attribute $\underset{\wedge}{Y}$-scores on a given testing occasion *t* is most naturally conceived as an interval of *s*'s life spanning minutes if not days; and even if we did artificially define the locus of $\underline{s}$'s-having-test-scores-$\underline{Y}$-on-occasion-$\underline{t}$ to be an instant of $\underline{s}$'s duration, we could not reasonably expect the dynamics of instantaneous $\underset{\wedge}{Y}$-scores to be usefully characterized by differential equations or even by difference equations for time displacements on the order of seconds or less. A great deal of practical importance lies within this abstractly obvious point, as will be touched upon in Chapter 4.

[[One reason for describing dynamic model (16/17) as "basic" is its meager manifest structure compared to the almost-unlimited potential for proliferating the same system variables under different translocators on both sides of the dynamic equation. Technically, (16) is a "first-order" auto-regression whose manifest (m+1)th-order extension is

$$(18) \qquad \text{In } \underline{D}/\overset{m}{\underline{f}},\ \underline{Yf}^{m+1} = \Phi(\underline{Yf}^{m},\underline{Zf}^{m},\underline{Yf}^{m-1},\underline{Zf}^{m-1},\ldots,\underline{Yf},\underline{Zf},\underline{Y},\underline{Z}) ,$$

albeit to appreciate the implicit scope of (16) it is worth mention that (18) can in turn be converted by translocation into a special case of form (16). With the possible exception of field theory in physics--about which, unhappily, I know very little--model (18) and its continuous-change counterpart is about as far as modern causal-systems theory has yet gone.[12] Yet many natural phenomena are not comfortably formalizable even within the generality of (18).

[[Consider, for example, the determination of height in animals by the heights of their parents. (Ignore that phenotypic parental height is not really itself a cause of offspring height, but is merely diagnostic of genetic factors that do the job.) Let $\underset{\wedge}{h}$ be $\underline{\text{Height}}$ over some suitably restricted subset $\underline{D}$ of temporal stages $\{\underline{s}\text{-at-}\underline{t}\}$ of bisexual organisms $\{\underline{s}\}$, while $\langle \underline{f}_{♂},\underline{f}_{♀}\rangle$ is the $\underline{\text{parents-of}}$ translocator defined previously (p. 26), i.e. $\underline{f}_{♂}(\underline{s}\text{-at-}\underline{t}) = \underline{s}$'s-father-at-time-

---

[12] The mathematical literature on system trajectories contains other models quite different from (18), notably spectral decompositions in time-series analysis, that likewise search for regularities in system growth; but these have little if anything to say about causality.

of-$\underline{s}$'s-conception and similarly for maternal translocator $\underline{f}_{\female}$. Then we may suppose

(19) For all $\underline{p}$ (= $\underline{s}$-at-$\underline{t}$) in $\underline{D}$, $\underset{\wedge}{h}(\underline{p}) = \phi(\underset{\wedge}{h}\underline{f}_{\male}(\underline{p}), \underset{\wedge}{h}\underline{f}_{\female}(\underline{p}), \underset{\wedge}{E}(\underline{p}))$,

or simply

(19') In $\underline{D}$, $\underline{h} = \phi(\underline{hf}_{\male}, \underline{hf}_{\female}, \underline{E})$,

where $\underset{\wedge}{E}$ is a tuple of specified and/or residual height sources additional to the two parental-height variables. If we want to be specific about the transducer in (19), we can further assume $\phi(\_\_,\_\_,\_\_) = g_1^{-1}(.5g_{1}\_\_ + .5g_{1}\_\_ + g_{2}\_\_)$ where g_1 scales Height as Height-in-inches and g_2 is some suitable numerical scaling of $\underset{\wedge}{E}$. This model illustrates first of all that auto-regressive locus displacements need not be just within-subject time differences. But its main point is that the manifest excursor in (19) runs backward: Instead of telling where certain causes have a certain kind of effect, as in (11/16/18), the translocators in (19) map the loci of its accounted-for effects into the loci of their causes. And the sequence of product-laws entailed by iteration of (19) is a precession, not a succession like (14). Thus from (19) we have

$$\text{In } \underline{D}/\underline{f}_{\male} \cap \underline{D}/\underline{f}_{\female}, \quad [\underline{hf}_{\male}, \underline{hf}_{\female}] = [\phi(\underline{hf}_{\male}^{2}, \underline{hf}_{\female}\underline{f}_{\male}, \underline{Ef}_{\male}), \phi(\underline{hf}_{\male}\underline{f}_{\female}, \underline{hf}_{\female}^{2}, \underline{Ef}_{\female})] ,$$

which integrates with (19') to give $\underset{\wedge}{h}$ as a function of grandparental heights together with supplementary sources $[\underset{\wedge}{E}\underline{f}_{\male}, \underset{\wedge}{E}\underline{f}_{\female}, \underset{\wedge}{E}]$ in domain $\underline{D}/\underline{f}_{\male} \cap \underline{D}/\underline{f}_{\female}$ of $\underline{D}$-members having both parents in $\underline{D}$.

[[Although model (19) can be reformalized in various ways to have a domain and manifest locus structure different from (19), I can find none of form (18) that can be repeatedly stepped forward to yield a form-(15) auto-regressive system with nonempty domain. Clearly, recursive causal processes whose single-step dynamics have a manifest tree structure of this sort can step forward indefinitely; yet to my knowledge, extant abstract systems theory has not yet worked out effective formalisms for iterating forward-branching systems. More generally, I submit that progress in our understanding of complex natural systems depends greatly upon our enhancing the articulate flexibility with which we conceive of locus structures.]]

Causal Metaprinciples.

Scientific practice exhibits many ways to explain laws, causal and otherwise. I shall briefly describe the most elemental of these, without detailing how they instantiate generic explanation schema (2), under a handy label for each. I characterize these as "elemental" because they often--some almost always--occur in combination as parts of more complicated explanations. The grade of causality in laws so explained is prevailingly problematic, even though intuition is not likely to dispute that we want to consider most of them "causal" in at least some superveniently weak sense. Indeed, it is precisely the prospect of finding explanations such as these for molar laws that commonsense insists are causal that mandates search for a coherent theory of grades or levels of causality.

Mediated Composition. If $\langle \underline{L}_1,\ldots,\underline{L}_m \rangle$ is a c-series (cf. p.69) of consecutively integrable laws whose product is $\underline{L}^*_{1m}$, then $\underline{L}^*_{1m}$ is a law due to $\underline{L}_1,\ldots,\underline{L}_m$ that is moreover causal if the latter are causal and have a suitably standard mediation structure. This paradigm for explaining wide-arched laws by identifying process mechanisms for their production has overwhelmingly dominated past accounts of this matter. In particular, factoring a received input/output dependency as a product of mediating causal operations has been classically the target of "black box" analysis in systems theory. And for psychologists of an older but still surviving generation, this was the model popularized by Tolman (1936) for unraveling the intricacies of behavior's overt lawfulness by appeal to "intervening variables" as unobserved causal mediators. Although there is no reason to hold the grade of causality in product-law $\underline{L}^*_{1m}$ inferior to that of its factors $\underline{L}_1,\ldots,\underline{L}_m$, $\underline{L}^*_{1m}$ is less fundamental than the latter not merely in being explanatorily derivative from them but also in usually having much narrower scope (i.e. breadth of domain) than any one of $\underline{L}_1,\ldots,\underline{L}_m$. Accordingly, to move beyond $\underline{L}^*_{1m}$ to disclosure of some parsing $\langle \underline{L}_1,\ldots,\underline{L}_m \rangle$ of the process mechanism generating $\underline{L}^*_{1m}$ is often a considerable advance in scientific knowledge.

Domain Constriction, weak and strong. For any causal law

L^+: In D^+, $y = \phi(X)$

and any proper subset D of D^+,

L: In D, $y = \phi(X)$

is also a causal law that is due jointly to L^+ and domain relation $D \subset D^+$. Formally this seems trivial; yet it is far from trivial for empirical research to establish L first and only afterward learn that specification of original domain D contains irrelevancies whose deletion expands L into L^+. Moreover, when D is so restricted that some component variables in $\underset{\wedge}{X}$ are essentially constant throughout D, what we are likely to have identified is not L but only

L^o: In D, $y = \phi'(X')$,

wherein $\underset{\wedge}{X}'$ comprises just the components of $\underset{\wedge}{X}$ that have nonzero variance in D, and ϕ' is the function on the range of $\underset{\wedge}{X}'$ derived by constriction of ϕ to the particular constancies in D on $\underset{\wedge}{X}$-components not in $\underset{\wedge}{X}'$. That is, if $\underset{\wedge}{X} = [\underset{\wedge}{X}', \underset{\wedge}{X}'']$ while X''_c is the constant value of $\underset{\wedge}{X}''$ in D, $\phi'(\_\_) = \phi(\_\_, X''_c)$.

Example. In pencil-shadows law L_{ps} (p. 45, above), numerical parameters 96 and 8 in L_{ps}'s transducer are due to constancies in domain D_{ps} of the pencil's height and positioning on the table. The contribution of these latter variables to a broader law L^+_{ps} of pencil shadows from which L_{ps} is derivative can readily be identified so long as the scope of our expansion is modest. Specifically, if D^+_{ps} is a broadening of D_{ps} that retains all the essential constancies in D^+_{ps} except for allowing variation in $\underset{\wedge}{h}$: *Pencil-height-in-inches* and $\underset{\wedge}{w}$: *Pencil-distance-in-inches-from-point-on-table-closest-to-light*, we have

L^+_{ps}: In D^+_{ps}, $y = \dfrac{h \cdot w}{x - h} + e$,

with $\underset{\wedge}{e}$ a negligible residual as before. Then L_{ps} derives from L^+_{ps} by selecting

$\underline{D}_{ps}$ to comprise just those $\underline{D}^{+}_{ps}$-objects whose values of $\underset{\wedge}{h}$ and $\underset{\wedge}{w}$ are 8 and 12, respectively, and constricting $\underline{L}^{+}_{ps}$'s transducer accordingly. Note that $\underline{L}^{+}_{ps}$ remains true if $\underline{D}^{+}_{ps}$ is replaced by its subset $\underline{D}_{ps}$, whereas $\underline{L}_{ps}$ does not generalize to $\underline{D}^{+}_{ps}$.

More generally, $\underset{\wedge}{X}'$ in an $\underline{L}^{o}$ so derived from $\underline{L}^{+}$ can be any subtuple of $\underset{\wedge}{X}$ that includes all $\underset{\wedge}{X}$-components having nonzero variance in $\underline{D}$. $\underline{L}^{o}$, too, is a causal law explained by domain constriction of $\underline{L}^{+}$; but recovery of $\underline{L}^{+}$ from $\underline{L}^{o}$ is much harder than from $\underline{L}$ insomuch as the former requires not only elimination of domain irrelevancies but also discovery of what locally constant variables make what difference for what parameters in transducer ϕ' of $\underline{L}^{o}$. (That a law's transducer can embody causal contributions from local constancies points out that the nature of causal transduction and the role of a law's domain preconditions in production of its output deserve far more discussion than I have given them here.) The difference between $\underline{L}$ and $\underline{L}^{o}$ is methodologically quite important; so to label this distinction we can refer to the transducer-narrowing move from $\underline{L}^{+}$ to $\underline{L}^{o}$ as strong domain constriction in contrast to weak domain restriction that retains the broader-scope transducer. That is, strong domain constriction may be viewed as weak domain constriction of $\underline{L}^{+}$ to $\underline{L}$, followed by the shift from $\underline{L}$ to $\underline{L}^{o}$ which, if wanting its own label, might be called "transducer absorption."

[For simplicity, I have identified Strong Domain Constriction as deriving $\underline{L}^{o}$ from $\underline{L}$ through transducer absorption of variables $\underset{\wedge}{X}''$ that are constant at some value $\underline{X}''_{c}$ in $\underline{D}$. However, let us say that subtuple $\underset{\wedge}{X}''$ of $\underset{\wedge}{X} = [\underset{\wedge}{X}', \underset{\wedge}{X}'']$ is quasi-constant at $\underline{X}''_{c}$ in $\underline{L}$ iff, for all values $\underline{X}'$ of $\underset{\wedge}{X}'$ and $\underline{X}''$ of $\underset{\wedge}{X}''$ whose joint occurrence in $\underline{D}$ is nomically possible, $\phi(\underline{X}', \underline{X}'') = \phi(\underline{X}', \underline{X}''_{c})$. That is, $\underset{\wedge}{X}''$ is quasi-constant in $\underline{L}$ iff the variance of $\underset{\wedge}{X}''$ allowed in $\underline{L}$'s domain is unable to produce any variance in $\underline{L}$'s output. Then everything said above about Strong Domain Constriction goes through as before if 'constant' is weakened to 'quasi-constant' and 'have nonzero variance' to 'lack quasi-constancy'. This point has passing importance in Chapter 5.]

Domain Translocation. For any non-null subset D'_f of the domain of any translocation function f, if f maps each member of D'_f into some object in the domain D_1 of causal law

$$L_1: \qquad \text{In } D_1,\quad y = \phi(X)\ ,$$

then

$$L_{1f}: \qquad \text{In } D'_f,\quad yf = \phi(Xf)$$

is a causal law for determination of $[\underset{\wedge}{y}f]$ by $[\underset{\wedge}{X}f]$ in D'_f that in general is due to L_1 together with the t-derivative constitutions of L_{1f}'s variables. (For clarification of L_{1f}'s notation, recall fn. 9a, p. 71a.)

> **Example**. Suppose that D_1 comprises (momentary stages of) humans, that $\underset{\wedge}{y}$ is Weight, and that $\underset{\wedge}{X}$ is a compound of physiological and dietary factors. Then L_1 says that the weight of any human o is caused under transduction ϕ by o's $\underset{\wedge}{X}$-status. If f is now the *wife-of* translocator over the domain D_f of monogamous males and, more restrictively than necessary, we take D'_f to consist of monogamous Englishmen, each o in D'_f has exactly one wife in the domain of L_1, and L_{1f} says that the wife-weight of any monogamous Englishman o is caused ϕ-wise by o's wifely $\underset{\wedge}{X}$-ness.

When L_{1m} is explained under Mediated Composition as the product of $\langle L_1,\ldots,L_m\rangle$, almost certainly some of the latter derive from more basic laws by Domain Translocation (cf. auto-regressive unfolding, p. 73f.). Prima facie, L_{1f} is just a domain constriction of L_1; yet conceptually, L_{1f}'s variables appear distinct from those of L_1, nor is D'_f always included in D_1. And if the manner in which variables $\underset{\wedge}{X}f$ and $\underset{\wedge}{y}f$ are initially identified by science Σ does not make manifest that these are in fact t-derivative from $\underset{\wedge}{X}$ and $\underset{\wedge}{y}$, respectively, say because what Σ has established is not L_{1f} as such but a less perspicuous conception

$$L''_1: \qquad \text{In } D''_1,\quad y'' = \phi(X'')$$

of the same factual generality signified by $\underline{L}_{1f}$, discovery that $\underline{D}''_1$, $\underset{\wedge}{X}''$, and $\underset{\wedge}{y}''$ are in fact $\underline{D}'_f$, $\underset{\wedge}{X}\underline{f}$, and $\underset{\wedge}{y}\underline{f}$, respectively, so that $\underline{L}''_1$ reduces by substitution of identicals to $\underline{L}_{1f}$ and is hence explained by $\underline{L}_1$, would be no small achievement for science Σ.

⟦*Note 1*. When $\underline{L}''_1$ is so reduced to $\underline{L}_{1f}$, Σ must also acknowledge the principles

$\underline{P}_1$: For each $\underline{o}_i$ in $\underline{D}''_1$ (= $\underline{D}'_f$), there is exactly one $\underline{o}_j$ in $\underline{D}_1$ such that $\underline{o}_j = \underline{f}(\underline{o}_i)$,
$\underline{P}_2$: For each $\underline{o}_i$ in $\underline{D}'_f$ and $\underline{o}_j$ in $\underline{D}_1$ such that $\underline{o}_j = \underline{f}(\underline{o}_i)$, $[\underset{\wedge}{X}'',\underset{\wedge}{y}''](\underline{o}_i) = [\underset{\wedge}{X},\underset{\wedge}{y}](\underline{o}_j)$.

$\underline{P}_2$ is not only a functional law of generic form (9), it is explanatory in that for each $\underline{o}$ in $\underline{D}'_f$ and $\underline{f}(\underline{o})$ in $\underline{D}_1$, $\underline{o}$ has its particular value of $[\underset{\wedge}{X}'',\underset{\wedge}{y}'']$ (= $[\underset{\wedge}{X}\underline{f},\underset{\wedge}{y}\underline{f}]$) *because* $\underline{f}(\underline{o})$ has whatever $[\underset{\wedge}{X},\underset{\wedge}{y}]$-state it does. But $\underline{P}_2$ is not a *causal* law; rather, it expresses a constitutional dependency of $\ulcorner[\underset{\wedge}{X}'',\underset{\wedge}{y}''];\underline{o}\urcorner$ upon $\ulcorner[\underset{\wedge}{X},\underset{\wedge}{y}];\underline{f}(\underline{o})\urcorner$ disclosed by analysis--an ontological, not conceptual, analysis--of the compositions of these events. It is hard to find a name for laws of this sort having just the right heuristic flavor, but "analytic reduction" or "supervenience analysis" come close. As for $\underline{P}_1$, whether this has explanatory force depends on details of the particular case--it may, for example, be true by definitional artifice (e.g., starting with a relation γ_f on $\underline{D}_f \times \underline{D}_1$ that is not a function, we might stipulate that $\underline{D}'_f$ comprises just the objects in $\underline{D}_f$ having exactly one γ_f-relatum in $\underline{D}_1$). But formally, $\underline{P}_1$ illustrates principles governing particular facts of locus structure that for at least some $\underline{f}$ (notably, ones that are genuine causal excursa) are explanatorily prior to any of the world's causal events.

⟦*Note 2*. Statements about t-derivative variables and their principles of explanation catch us in an awkward ontological bind whose optimal philosophic management is still unclear. Consider the sentence, 'Spratt's wife is fat.' We can parse this as attributing either the property *Having-a-fat-wife* to Spratt, or *Fatness* to Spratt's wife--which is to say that this sentence's gerundive, 'Spratt's wife's being fat', seems ambiguous between naming an event whose locus is Spratt and one whose locus is Spratt's wife. More generally, when variable $[\underset{\wedge}{x}\underline{f}]$ is t-derivative from variable $\underset{\wedge}{x}$, we can apparently construe gerundized true sentence '$\underset{\wedge}{x}\underline{f}(\underline{o}) = \underline{x}$' to signify either the event $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$ or the manifestly quite different event $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}\urcorner$. Now, an ontology that takes such translocational multicopying of x-events seriously, to the point of distinguishing $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}\urcorner$ from $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$ causally, seems surrealistic at best. Yet for reasons deeply rooted in our most basic machinery of thought, it is extraordinarily difficult to abandon talking *as though* $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}\urcorner$ is distinct from $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$. That is, we have no ready way to identify Spratt's having a fat wife with Mrs. Spratt's being fat. Someday we will surely arrive at an understanding of semantics that can translate such hopefully harmless fables into ontological austerities that need no apologies for pretense. But until then, we must try to imagine coherently separate positions for $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$ and $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}\urcorner$ in our models of the world's explanatory order. It is straightforward enough to hold $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$, together with the fact that $\underline{o}$ has just one $\underline{f}$-relatum, analytically (*not* causally) responsible for $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}\urcorner$. But if we hope to envision causal laws in which $[\underset{\wedge}{x}\underline{f}]$ is an input variable, the productivity force we attribute to $[\underset{\wedge}{x}\underline{f}]$ can only be a supervenient shadow of $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o})\urcorner$'s grade of causality. To illustrate, suppose that $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o}_i)\urcorner$ is a cause of some $\ulcorner\underset{\wedge}{y};\underline{f}(\underline{o}_j)\urcorner$ and hence t-derivatively of $\ulcorner\underset{\wedge}{y}\underline{f};\underline{o}_j\urcorner$. If we wish to say that $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}_i\urcorner$, too, is a cause of $\ulcorner\underset{\wedge}{y};\underline{f}(\underline{o}_j)\urcorner$ or of $\ulcorner\underset{\wedge}{y}\underline{f};\underline{o}_j\urcorner$, how do we put $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o}_i)\urcorner$ and $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}_i\urcorner$ together in a single story about production of $\ulcorner\underset{\wedge}{y};\underline{f}(\underline{o}_j)\urcorner$ and $\ulcorner\underset{\wedge}{y}\underline{f};\underline{o}_j\urcorner$? Our best move is *not* to try combining these in the way we connect two events that work together, conjointly or one through the other's mediation, in bringing about a shared effect. Instead, we need to segregate $\ulcorner\underset{\wedge}{x}\underline{f};\underline{o}_i\urcorner$ from $\ulcorner\underset{\wedge}{x};\underline{f}(\underline{o}_i)\urcorner$ in non-interacting stories at different levels of causality. So in particular, when the law $\underline{L}_{1f}$ derived by translocation from causal law $\underline{L}_1$ is parsed as attributing causal linkages to events whose loci in $\underline{D}'_f$ are artificial displacements of real causal loci that are in or have been previously translocated into $\underline{D}_1$, the grade of causality involved in $\underline{L}_{1f}$ must be superveniently derivative from, not the same as, that in $\underline{L}_1$.

[Note 3. Whether we can disambiguate this metaprinciple's explanatory-order claim, that $\underline{L}_{1f}$ is "in principle" due to $\underline{L}_1$, as a strict no-exceptions rather than just an expectation in the main depends on whether $\underset{\wedge}{y}$ is identical with $[\underset{\wedge}{y}\underline{ff}^{-1}]$ whenever the restriction of translocator $\underline{f}$ to its subdomain $\underline{D}'_f$ has an inverse $\underline{f}^{-1}$, or whether there is instead some sense in which $[\underset{\wedge}{y}\underline{ff}^{-1}]$ is superveniently distinct from $\underset{\wedge}{y}$ despite $\underline{ff}^{-1}$'s being simply an Identity function on $\underline{D}_1$. (Fn. 4 in Chapter 1 points out why the latter suggestion is not quite so ludicrous as it seems.) For example, if Wife's-husband's-weight is the variable whose value for any monogamous Englishman is the weight of his wife's husband, is Spratt's-having-a-wife's-husband's-weight-of-163-lbs. the very same event as Spratt's-weighing-163-lbs., as commonsense insists, or does the logic by which we work t-derivative events into the world's becausal order require these to be distinguished? Unless $[\underset{\wedge}{y}\underline{ff}^{-1}]$ generally differs from $\underset{\wedge}{y}$, however, a pair of laws $\langle \underline{L}_1, \underline{L}_{1f} \rangle$ instantiating the Domain Translocation schema may give explanatory priority to $\underline{L}_{1f}$ rather than to $\underline{L}_1$--as you can see by replacing $\langle \underset{\wedge}{y}, \underset{\wedge}{X} \rangle$ in $\underline{L}_1$ by $\langle \underset{\wedge}{y}\underline{f}, \underset{\wedge}{X}\underline{f} \rangle$ and both occurrences of $\underline{f}$ in $\underline{L}_{1f}$ by $\underline{ff}^{-1}$.]

Output Abstraction. If variable $\underset{\wedge}{z}$ is a-derivative from the output $\underset{\wedge}{Y}$ of a causal law

$$\underline{L}_Y: \qquad \text{In } \underline{D},\ \ \underline{Y} = \Phi(\underline{X}) \ ,$$

i.e., if $\underset{\wedge}{z} = [g\underset{\wedge}{Y}]$ for some abstractor function g, then

$$\underline{L}_{gY}: \qquad \text{In } \underline{D},\ \ \underline{z} = g\,\Phi(\underline{X}) \qquad (\ \underset{\wedge}{z} = [g\underset{\wedge}{Y}]\ )$$

is a causal law, due to $\underline{L}_Y$ and $\underset{\wedge}{z}$'s constitutional derivation from $\underset{\wedge}{Y}$, whose transducer is the composition of Φ into g and whose grade of causality is presumably more rarified than that of $\underline{L}_Y$. For example, suppose that $\underset{\wedge}{Y}$ is the singleton *Height* variable over bisexual organisms, $\underset{\wedge}{X}$ is a tuple of Height sources, and $\underset{\wedge}{z}$ is the binary variable of being/not-being at least one meter tall. Then an organism's $\underset{\wedge}{X}$-condition causally determines whether it is at least one meter tall by producing its value on the continuous height variable from whose range the binary height classification is an analytic abstraction. More generally, the metaprinciple of Output Abstraction says that any molar event $\ulcorner g\underset{\wedge}{Y};\underline{o}\urcorner$ is caused, perhaps superveniently, by any causal source of the compound event $\ulcorner \underset{\wedge}{Y};\underline{o}\urcorner$ from which $\ulcorner g\underset{\wedge}{Y};\underline{o}\urcorner$ is a molar abstraction.

[*Note*. By rights, this metaprinciple (and similarly for Input Abstraction, below) should also extend to cases wherein g in $\underline{L}_{gY}$ is a scaling function; and to be sure, laws encorporating scaled variables need to be positioned somewhere among the world's becausings. Yet if $\underline{L}_{gY}$ is due to $\underline{L}_Y$ when g is a scaling function, then $\underline{L}_Y$ is also due to $\underline{L}_{gY}$--in flagrant violation of the antisymmetry of explanation--unless we can disclaim that variable $g^{-1}g\underset{\wedge}{Y}$ is identical with variable $\underset{\wedge}{Y}$. As a last resort, we can support that disclaimer on grounds mentioned in fn. 4, p. 22; but surely there are better ways to cope with this difficulty. So let us sneak by with some evasive pleading that the theory of acausal explanation is still too nascent for all its major problems to demand immediate solution.]

<u>Input Abstraction</u>. Let

$\underline{L}_{yX}$: $\quad$ In $\underline{D}$, $\underline{y} = \phi g(\underline{X})$

be a causal law whose transducer can be factored as the composition of some function g into another, ϕ. That is, $\phi g(\_\_) = \phi(g(\_\_))$. And write $\underset{\wedge}{Z}$ for the simple or compound variable derived from $\underset{\wedge}{X}$ by viewing g in $\underline{L}_{yX}$ not as a transducer composant but as an abstractor applied to $\underset{\wedge}{X}$, i.e., $\underset{\wedge}{Z} = [g\underset{\wedge}{X}]$. ⟦We allow that some $\underset{\wedge}{X}$-components may persist unabstracted as components of $\underset{\wedge}{Z}$, as when, e.g., $\underset{\wedge}{X} =_{def} [\underset{\wedge}{X}_1, \underset{\wedge}{X}_2]$, $g(\underline{X}) =_{def} \langle g_1(\underline{X}_1), \underline{X}_2 \rangle$, and $\underset{\wedge}{Z} = [g\underset{\wedge}{X}] = [[g_1\underset{\wedge}{X}_1], \underset{\wedge}{X}_2]$. ⟧ Then

$\underline{L}_{yZ}$: $\quad$ In $\underline{D}$, $\underline{y} = \phi(\underline{Z})$ $\quad$ ($\underset{\wedge}{Z} = [g\underset{\wedge}{X}]$)

is a supervenient causal law that is due to $\underline{L}_{yX}$ together with $\underset{\wedge}{Z}$'s a-derivational nature. For example, if $\bar{\underline{h}}_0$ is the a-derivative variable <u>Mean-parental-height-in-inches</u> (see p. 26), we can infer from our previously (p. 80) conjectured law

$$\text{In } \underline{D},\quad \underline{h} = g_1^{-1}(.5g_1\underline{hf}_{♂} + .5g_1\underline{hf}_{♀} + g_2\underline{E})$$

of height inheritance that

$$\text{In } \underline{D},\quad \underline{h} = g_1^{-1}(\bar{\underline{h}}_0 + g_2\underline{E})$$

is also, at some level of molar abstraction, a causal law under which $\underline{o}$'s height is brought about by $\underline{o}$'s mean-parental-height together with supplementary input $\ulcorner \underset{\wedge}{E}; \underline{o} \urcorner$. In essence, what this metaprinciple claims is that when several different values of $\underset{\wedge}{X}$ have identical effects on $\underset{\wedge}{y}$, then the disjunctive property of having one of these $\underset{\wedge}{y}$-equivalent alternatives on $\underset{\wedge}{X}$ can itself be treated as a molar source of this same $\underset{\wedge}{y}$-effect--albeit not on the same causality level as the former, else we would have the problem of explaining how the disjunction and the $\underset{\wedge}{X}$-value which entails it for a particular $\_$ work conjointly in determination of $\ulcorner \underset{\wedge}{y};\_ \urcorner$. It is not plain that every L_{yZ} derived by input abstraction from a causal law is itself causal at any level of supervenience; and it may well turn out that severe additional qualifications are needed here. Even so, unless laws derived by input abstraction generally enjoy some grade of causal stature, attribution of causal efficacy to commonsense molar events could never be more than fatuity.

Output Compounding. Let

$$\underset{\sim}{L}: \qquad \left\{ \text{In } D_k, \ \ y_k = \phi_k(X) \right\} \qquad (k = 1,\ldots,m)$$

be an m-tuple of causal laws with the same input $\underset{\wedge}{X}$. If $\underset{\wedge}{Y} =_{\text{def}} [\underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_m]$, $\Phi =_{\text{def}} \langle \phi_1,\ldots,\phi_m \rangle$, and D^*_{1m} is the intersection of domains $D_1,\ldots,D_m$, then

$$L_{1m}: \qquad \text{In } D^*_{1m}, \ \ Y = \Phi(X)$$

is also a causal law due to laws $\underset{\sim}{L}$. This is one metaprinciple that seems entirely unproblematic, and with the aid of Input Expansion, below, is our basis for allowing the outputs of causal laws to be compound variables. The importance of combining laws $\underset{\sim}{L}$ into a single compound law L_{1m} is that L_{1m} thereby becomes available as a basis for further derivation, by Output Abstraction and Input Abstraction, of laws that account for complex molar events $\{\ulcorner g\underset{\wedge}{Y};a \urcorner\}$.

Input Expansion. To conjoin a given law

$$L_k: \qquad \text{In } D, \ \ y_k = \phi_k(X_k)$$

with others by Output Compounding, L_k's input may first require inflation by local irrelevancies to which its transducer gives null weight. Suppose that $\underset{\wedge}{X}_k$ is only a proper subtuple of the totality $\underset{\wedge}{X}$ of input dimensions in the laws to be conjoined. To expand L_k to receive all of $\underset{\wedge}{X}$ as input, let σ_k be the component-selector function, from the range of $\underset{\wedge}{X}$ into the range of $\underset{\wedge}{X}_k$, that maps each value X of

$\underset{\wedge}{X}$ into the $\underset{\wedge}{X}_k$-value $\underline{X}_k$ embedded in $\underline{X}$. Then as illustrated by $\underline{L}_e$ on p. 68,

$$\underline{L}_{k\sigma}: \qquad \text{In } \underline{D},\ \ \underline{y}_k = \phi_k\sigma_k(\underline{X})$$

is a law that is also causal if the components of $\underset{\wedge}{X}$ not in $\underset{\wedge}{X}_k$ have no effect on $\underset{\wedge}{y}_k$ in $\underline{D}$ that is not wholly mediated by $\underset{\wedge}{X}_k$, and indeed is more basic than $\underline{L}_k$ in that case. (Details of this mediation condition are developed in my "Complexities of mediation structure" document cited in fn. 8.) Transducer $\phi_k\sigma_k$ in $\underline{L}_{k\sigma}$ illustrates what is meant by saying that a law may assign "null weight" to some components of its input. For $\phi_k\sigma_k$ simply ignores values of $\underset{\wedge}{X}$-components that are not also in $\underset{\wedge}{X}_k$, i.e., if $\underline{X}_k$ is the value of $\underset{\wedge}{X}_k$ embedded in $\underset{\wedge}{X}$-value $\underline{X}$, $\phi_k\sigma_k(\underline{X}) = \phi_k(\underline{X}_k)$. Whenever we undertake to explain a molar event $\ulcorner g[\underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_m];\underline{o}\urcorner$ for some $\underline{o}$ in $\underline{D}$ by appeal to an array of molecular laws $\{\underline{y}_k = \phi_k(\underline{X}_k)\}$ ($\underline{k} = 1,\ldots,\underline{m}$) over $\underline{D}$, we must first extend each of the latter to have as common input all the different variables in $\langle\underset{\wedge}{X}_1,\ldots,\underset{\wedge}{X}_m\rangle$--which, depending on the mediation structure of these variables, may or may not be possible simply by Input Expansion--before Output Compounding and thereafter Output Abstraction can be applied.

Scientific systemacy 3V. Molar explanations.

It is evident from the foregoing that lawfulness derived from an ensemble of causal laws may well have dubious causal quality. But that should hold no surprise; for more generally, causal laws can generate unlimited patterns of acausal concordance, functional or otherwise.

Suppose, for example, that compound variable $\underset{\wedge}{Y}$ has causal origin

$$\underline{L}_Y: \qquad \text{In } \underline{D},\ \ \underline{Y} = \Phi(\underline{X}) \qquad (\ \underset{\wedge}{Y} = [\underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_n],\ \underset{\wedge}{X} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m],\ \Phi = \langle\phi_1,\ldots,\phi_n\rangle\ ).$$

As $\underset{\wedge}{X}$ takes on different values for assorted members of $\underline{D}$, the component variables in $\underset{\wedge}{Y}$ are not free to vary independently of one another but are constrained under the component functions in Φ by their common source $\underset{\wedge}{X}$. In ideal cases this can preclude occurrence in $\underline{D}$ of many logically possible score combinations on

$\langle \underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_n\rangle$. Indeed, if the number $\underline{m}$ of $\underset{\wedge}{X}$-components is smaller than the number $\underline{n}$ of components in $\underset{\wedge}{Y}$ while the first $\underline{m}$ component functions in Φ satisfy certain mild orthodoxies, we can write $\underset{\wedge}{Y} = [\underset{\wedge}{Y}_1, \underset{\wedge}{Y}_2]$ and $\Phi = \langle \Phi_1, \Phi_2 \rangle$ where Y_1 and Φ_1 contain the first $\underline{m}$ components of $\underset{\wedge}{Y}$ and Φ, respectively, and have that Φ_1 has an inverse.[14] Then insomuch as $\underline{Y}_1 = \Phi_1(\underline{X})$ and $\underline{Y}_2 = \Phi_2(\underline{X})$ in $\underline{D}$ (immediate from $\underline{L}_Y$), it follows that

$$\text{In } \underline{D}, \quad \underline{X} = \Phi_1^{-1}(\underline{Y}_1) \quad \text{and} \quad \underline{Y}_2 = \Phi_2\,\Phi_1^{-1}(\underline{Y}_1) \ .$$

Neither of these latter regularities expresses a causal dependency of $\underset{\wedge}{X}$ or $\underset{\wedge}{Y}_2$ upon $\underset{\wedge}{Y}_1$, but both are lawful enough to sustain genuine predictions based on $\underset{\wedge}{Y}_1$-information. That is, given knowledge of a compound causal law $\underline{L}_Y$ having the ideal properties just described--though how we might come by that knowledge if $\underset{\wedge}{X}$-scores are not directly observable remains to be seen--learning where some $\underline{o}$ in $\underline{D}$ stands just on subtuple $\underset{\wedge}{Y}_1$ of $\underset{\wedge}{Y}$ allows us to recover $\underline{o}$'s standing on $\underset{\wedge}{Y}_1$-source $\underset{\wedge}{X}$ and from there to predict what value $\underline{o}$ has on other effects $\underset{\wedge}{Y}_2$ of $\underset{\wedge}{X}$. And as bonus, we get to explain both the predicted event $\ulcorner \underset{\wedge}{Y}_2;\underline{o}\urcorner$ and our evidence $\ulcorner \underset{\wedge}{Y}_1;\underline{o}\urcorner$ for it as due to the inferred underlying event $\ulcorner \underset{\wedge}{X};\underline{o}\urcorner$.

To be sure, this predictive ideal is never entirely realistic. For noise factors (measurement error, unique sources, stochastic residuals, or whatever) on the input side of any multivariate causal law $\underline{L}_Y$ are always numerous enough to thwart errorless identification of $\ulcorner \underset{\wedge}{X};\underline{o}\urcorner$, or even certain exclusion of any possibility for $\underset{\wedge}{Y}_2(\underline{o})$ ($\underset{\wedge}{Y} = [\underset{\wedge}{Y}_1, \underset{\wedge}{Y}_2]$), from any part $\ulcorner \underset{\wedge}{Y}_1;\underline{o}\urcorner$ of $\ulcorner \underset{\wedge}{Y};\underline{o}\urcorner$. Even so, $\underline{L}_Y$ will still induce a distinctively patterned distribution of $\underset{\wedge}{Y}$-scores not only in $\underline{D}$ but also in observably finite subsets of $\underline{D}$. And so we arrive at the two great problems of raw macro-phenomena that drive the modern methodology of statistical

---

[14] More generally, without imposition of "orthodoxies" on $\langle \phi_1,\ldots,\phi_n\rangle$, if $\underline{n}$ is sufficiently large for fixed $\underline{m}$ there will likely be some integer $\underline{r} \le \underline{n}$ such that $\langle \phi_1, \ldots, \phi_r\rangle$ has an inverse. But $\underline{r}$ may well be larger than $\underline{m}$; and to insure the existence of this $\underline{r}$ even when $\underline{n}$ much exceeds $\underline{m}$ we need some constraints on $\langle \phi_1,\ldots,\phi_n\rangle$.

data analysis: Given that all components of compound variable $\underset{\wedge}{Y}$ in $\underline{L}_Y$ are observable but that some or all components of $\underset{\wedge}{X}$ are not, (*a*) how can we usefully characterize the $\underset{\wedge}{Y}$-score patterning in sample populations drawn from $\underline{D}$, and (*b*) how can we exploit this pattern information to infer decently reliable conclusions about the transducer $\underline{\Phi}$ and underlying $\underset{\wedge}{X}$-values from which our observed $\underset{\wedge}{Y}$-scores have arisen? The first of these is the problem of data *analysis*; the second, that of data *interpretation*.

In theory, the answer to (*a*) is straightforward: Every statistic defined by a-derivation on the range of $\underset{\wedge}{Y}$'s sample distributions (see p. 28, above) is a dimension of molar $\underset{\wedge}{Y}$-patterning; and although the totality of conceivable pattern dimensions is unworkably large (i.e. infinite), it only remains to choose which manageably small subset thereof seems most interpretively informative. In practice, of course, a great deal of sophisticated effort has gone into establishing traditions for [redacted] such choices, and many important technicalities remain open to further inquiry (including some pretty basic foundation gaps that need filling with something more solid than the pious presumptions that now patch them over.) Yet the essence of statistical data analysis/interpretation can be deftly summarized by our now-established SLese formalisms, and is well worth thinking through not so much for what it reveals about scientific knowledge as for its perspicuity as a paradigm of molar explanation.

Statistical aggregates.

In a typical application of modern statistics to empirical research, we make observations on a tuple $\underset{\wedge}{Y}$ of numerically scaled data variables for each individual object $\underline{o}$ in a size-$\underline{N}$ sample $\underline{s}_N$ of some population $\underline{D}$, while hypothesizing that $\underset{\wedge}{Y}$ is produced in $\underline{D}$ by numerically scaled but generally unobserved sources $\langle \underset{\wedge}{X}, \underset{\wedge}{E} \rangle$ (in which $\underset{\wedge}{E}$ is distinguished for later treatment as a tuple of "random" variables) according to a causal law

(20) $$\text{In } \underline{D},\quad \underline{Y} = \Phi_a(\underline{X},\underline{E})\ .$$

We assume that (20)'s transducer belongs to some identified class $\underset{\sim}{\Phi}$ of functions, but that its particular $\underline{a}$-indexed instance within this class is unknown. (In practice, '$\underline{a}$' demarks a finite tuple of algebraic parameters within a computable function-form $\underset{\sim}{\Phi}$.) We can choose form $\underset{\sim}{\Phi}$ to equate some or all components of $\underset{\wedge}{X}$ with particular components of $\underset{\wedge}{Y}$; so some of the nominally unobserved inputs in (20) can in fact be data variables. ⟦Also, source variables $\langle X,E\rangle$ can be under some array of conjectured constraints expressible e.g. in form

(20-add) $$\text{In } \underline{D},\quad \Psi_b(\underline{X},\underline{E}) = 0\ ,$$

i.e., $\langle \psi_{b1}(\underline{X},\underline{E}) = 0,\ \psi_{b2}(\underline{X},\underline{E}) = 0,\ \ldots\rangle$, in which Ψ_b is a function selected by unknown parameter tuple $\underline{b}$ from known form $\underset{\sim}{\Psi}$. In particular, (20-add) may be an algebraic reorganization of a recursive system $\{\text{In } \underline{D},\ \underline{x}_j = \psi''_{bj}(\underline{X}''_j,\underline{E}''_j)\}$ of causal laws wherein $\underset{\wedge}{x}_j$ is some component of $\underset{\wedge}{X}$ while $\underset{\wedge}{X}''_j$ and $\underset{\wedge}{E}''_j$ are subtuples of $\underset{\wedge}{X}$ and $\underset{\wedge}{E}$, respectively.⟧

The most familiar version of hypothesis schema (20), one that research practice has exploited prodigiously, is the basic "general linear model" under which $\underset{\wedge}{X} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m]$, $\underset{\wedge}{Y} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m,\underset{\wedge}{y}]$, $\underset{\wedge}{E} = [\underset{\wedge}{e}]$, and the ($\underline{m}$+1)th component ϕ_a of $\Phi_a = \langle\phi_1,\ldots,\phi_m,\phi_a\rangle$ is a linear function of $[\underset{\wedge}{X},\underset{\wedge}{e}]$ with coefficients $\underline{a} = \langle\underline{a}_0,\underline{a}_1,\ldots,\underline{a}_m\rangle$ and 1. (Components $\phi_1,\ldots,\phi_m$ of Φ_a trivially identify each $\underset{\wedge}{x}_i$ on the left with $\underset{\wedge}{x}_i$ on the right.) Then the model's nontrivial part is just

(20.1) $$\text{In } \underline{D},\quad \underline{y} = \underline{a}_0 + \underline{a}_1\underline{x}_1 + \ldots + \underline{a}_m\underline{x}_m + \underline{e}\ \ ,$$

in which scores on $\underset{\wedge}{y},\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m$ for sampled $\underline{D}$-members are known but coefficients $\underline{a}$ and scores on random residual $\underset{\wedge}{e}$ are not. Or if we waive our model-standardizing but otherwise needless constraint that all data variables in (20) are components of output compound $\underset{\wedge}{Y}$, we can say more simply that the basic General Linear Model is (20)'s instance wherein $\underset{\wedge}{Y}$ and $\underset{\wedge}{E}$ are singletons and (20) has linear form (20.1).

As is well known, (20.1) in turn subsumes group comparisons, Analysis of Variance/Covariance, and Regression Analysis with any parametrically linear form of regression surface. Another major instance of generic model (20) is Factor Analysis, wherein $\underset{\wedge}{X}$ is a tuple of unobserved source variables postulated to explain the sample covariances observed among data variables $\underset{\wedge}{Y}$ while $\underset{\wedge}{E}$ comprises unobserved "unique" residuals of these outputs. And in certain very recent advances of multivariate analysis, each $\underline{D}$-member is an enduring subject for which the value of each component $\underset{\wedge}{y}_j$ of $\underset{\wedge}{Y}$ is a "repeated measurements" sequence of scores on some variable $\underset{\wedge}{y}'_j$ at different times in that subject's developmental history.

To develop sampling theory for recovery of unknowns in (20), let $\underline{D}^*_N$ comprise all $\underline{N}$-tuples $\underline{s}_N$ of individuals from $\underline{D}$ such that $\underline{s}_N$ satisfies certain co-selection constraints of obscure factual character but denoted by such statisticians' phrases as "independent random sample." And for $\underline{k} = 1,\ldots,\underline{N}$, let $\underline{f}_k$ be the translocator from $\underline{D}^*_N$ into $\underline{D}$ such that each $\underline{f}_k(\underline{s}_N)$ is the $\underline{k}$th member of sample $\underline{s}_N$. Then by Domain Translocation, (20) generates

$$(21) \qquad \text{In } \underline{D}^*_N, \quad \left\{ \underline{Yf}_k = \Phi_a(\underline{Xf}_k, \underline{Ef}_k) \right\} \qquad (\, \underline{k} = 1,\ldots,\underline{N}\,) \quad ,$$

and from there, by Input Expansion and Output Compounding,

$$(22) \qquad \text{In } \underline{D}^*_N, \quad \underline{Y}^* = \Phi^*_a(\underline{X}^*, \underline{E}^*)$$

in which

$$\underset{\wedge}{Y}^* =_{\text{def}} [\underset{\wedge}{Y}\underline{f}_1,\ldots,\underset{\wedge}{Y}\underline{f}_N], \quad \underset{\wedge}{X}^* =_{\text{def}} [\underset{\wedge}{X}\underline{f}_1,\ldots,\underset{\wedge}{X}\underline{f}_N], \quad \underset{\wedge}{E}^* =_{\text{def}} [\underset{\wedge}{E}\underline{f}_1,\ldots,\underset{\wedge}{E}\underline{f}_N],$$

and

$$\Phi^*_a =_{\text{def}} \langle \Phi_a\sigma_1, \ldots, \Phi_a\sigma_N \rangle$$

where each σ_k is the component-selector function that maps any value of $[\underset{\wedge}{X}^*, \underset{\wedge}{E}^*]$ into the value of $[\underset{\wedge}{X}\underline{f}_k, \underset{\wedge}{E}\underline{f}_k]$ embedded therein. In practice, values of $\underset{\wedge}{Y}^*$ (and similarly for $\underset{\wedge}{X}^*$ and $\underset{\wedge}{Z}^*$) are often written as Variables-by-Individuals score matrices, i.e., $\underset{\wedge}{Y}^*(\underline{s}_N) = \{y_{jk}\}$ wherein y_{jk} is the score on the $\underline{j}$th $\underset{\wedge}{Y}$-variable for the *k*th member of sample s_N. This matrix reading of Y^*, X^*, and E^* is henceforth presumed here. Then Φ^*_a in (22) is the function that maps each score-matrix pair $\langle X^*, E^*\rangle$ into the score matrix whose *k*th column ($k = 1,2,\ldots,N$) is the Φ_a-transform of the *k*th columns of X^* and E^*.

[When model (20) includes additional source-variable constraints (20-add), these are similarly compounded into

(22-add) $$\text{In } D^*_N,\quad \Psi^*_b(X^*,E^*) = 0 \quad (\Psi^*_b =_{\text{def}} \langle \psi_b\sigma_1, \ldots, \psi_b\sigma_N\rangle)\ .]$$

To continue illustration with the General Linear Model, when (20) is specialized to (20.1), (22) becomes in the notation of matrix algebra just

(22.1) $$\begin{aligned}\text{In } D^*_N,\quad y^* &= [a_0\ a_1\ \ldots\ a_m\ 1]\begin{bmatrix}1^*\\ X^*\\ e^*\end{bmatrix}\\ &= a_0 1^* + a_1 x^*_1 + \ldots + a_m x^*_m + e^*\ ,\end{aligned}$$

wherein 1^* is an *N*-tuple of 1s, and each of $y^*, x^*_1, \ldots, x^*_m, e^*$ is the *N*-tuple (row vector) of scores in s_N on the variable indicated. Although all scores on $\underset{\wedge}{e}$ and $\underset{\wedge}{X}$ for all sample members are in the righthand side of (22.1), this equation's algebra selects just the *k*th components of e^* and the rows of X^* for determining the *k*th component of y^*. That is, with far less algebraic dexterity but greater notational resemblance to generic formula (22), (22.1) can be equivalently written as

(22.1*a*) $$\begin{aligned}\text{In } D^*_N,\quad \langle \ldots, y_k, \ldots\rangle = \langle \ldots,\ & a_0 + a_1\sigma_k(x_{11},\ldots,x_{1N}) + \ldots + a_1\sigma_k(x_{m1},\\ &\ldots, x_{mN}) + \sigma_k(e_1,\ldots,e_N),\ \ldots\rangle\ .\end{aligned}$$

In (22.1*a*), y_k, x_{hj}, and e_j are respectively the *k*th element of vector y^* of scores on $\underset{\wedge}{y}$, the *hj*th element of matrix X^* of scores on $\underset{\wedge}{X}$, and the *j*th element of vector e^* of scores on $\underset{\wedge}{e}$ for an arbitrary *N*-tuple of *D*-members; and component-selector function σ_k picks out the *k*th element of any *N*-tuple to which it is applied.

Let us pause to flesh out the meaning of formalism (22). We have begun with a compound $\underset{\wedge}{Y} = [\underset{\wedge}{y}_1, \ldots, \underset{\wedge}{y}_n]$ of numerically scaled variables whose values we have actually learned for certain objects in our vicinity. And for reasons seldom closely scrutinized in practice, we feel it worthwhile to conjecture that for each member $\underline{o}$ of some domain $\underline{D}$ decently sampled by the individuals whose $\underset{\wedge}{Y}$-scores we have observed, $\underline{o}$'s score-tuple $\langle \underset{\wedge}{y}_1(\underline{o}),\ldots,\underset{\wedge}{y}_n(\underline{o})\rangle$ on $\underset{\wedge}{Y}$ is caused by $\underline{o}$'s scores $\langle \underset{\wedge}{X}(\underline{o}),\underset{\wedge}{E}(\underline{o})\rangle$ on numerically scaled compound source-variables $\underset{\wedge}{X} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m]$ and residual $\underset{\wedge}{E} = [\underset{\wedge}{e}_1,\ldots,\underset{\wedge}{e}_n]$ under a transducer $\Phi_a = \langle \phi_{a1},\ldots,\phi_{an}\rangle$ (i.e. $\underline{y}_j = \phi_{aj}(\underline{X},\underline{E})$ for each $\underline{j} = 1,\ldots,\underline{n}$) specified by still-unknown parameters $\underline{a} = \langle\ldots,\underline{a}_h,\ldots\rangle$ in a stipulated function-form $\underset{\sim}{\Phi}$. For example, $\underset{\wedge}{y}_1(\underline{o}),\ldots,\underset{\wedge}{y}_n(\underline{o})$ might be person $\underline{o}$'s scores on an array of small intellectual tasks (test problems) on which we think performance in a certain testing situation partly definitive of $\underline{D}$ is due mainly to the performer's standing on an array $\underset{\wedge}{X}$ of ability/personality/motivational factors together with assorted chance disturbances that can be aggregated into one residual $\underset{\wedge}{e}_j$ for each performance dimension $\underset{\wedge}{y}_j$. (It is not obligatory to allocate exactly one residual to each component of output compound $\underset{\wedge}{Y}$, but that is standard in practice.) It should be noted that model (20) for single $\underline{D}$-members is itself explicitly or implicitly derived from $\underline{m}$ sub-models

(20') $$\text{In } \underline{D},\ \{\underline{y}_j = \phi'_j(\underline{x}'_{j1},\ldots,\underline{x}'_{jm_j},\underline{e}_j)\} \qquad (\underline{j} = 1,\ldots,\underline{n}),$$

one for each $\underset{\wedge}{Y}$-dimension $\underset{\wedge}{y}_j$, wherein $\langle \underset{\wedge}{x}'_{j1},\ldots,\underset{\wedge}{x}'_{jm_j}\rangle$ is a not-necessarily-proper subtuple of $\underset{\wedge}{X}$, and $\underset{\wedge}{e}_j$ is the $\underline{j}$th component of $\underset{\wedge}{E}$. The righthand side of each equation in (20') is first converted to $\phi_j(\underline{x}_1,\ldots,\underline{x}_m,\underline{e}_1,\ldots,\underline{e}_n)$, i.e. $\phi_j(\underline{X},\underline{E})$, by Input Expansion under presumption that any components of $[\underset{\wedge}{X},\underset{\wedge}{E}]$ not in $[\underset{\wedge}{x}'_{j1},\ldots,\underset{\wedge}{x}'_{jm_j},\underset{\wedge}{e}_j]$ have no effect on $\underset{\wedge}{y}_j$ unmediated by the latter, and those expansions are then conjoined by Output Compounding to yield (20). Until recently, applied data analysis seldom started with component models (20') in which the non-residual input to any one $\underset{\wedge}{y}_j$ was less than total input array $\underset{\wedge}{X}$. But that is no longer true of modern recursive-modelling practices.

What we initially surmise about the $\underset{\wedge}{Y}$-sources $\underset{\wedge}{X}$ in (20) varies greatly from one application to another. Often we have views on the substantive nature of some

$\underset{\wedge}{X}$-dimensions $\underset{\wedge}{x}_j$, e.g. theories of intelligence and personality, even when $\underset{\wedge}{x}_j$-scores are not directly observable. And sometimes we not merely know what $\underset{\wedge}{x}_j$-ness *is*, but believe it to be measured by one of our $\underset{\wedge}{Y}$-variables $\underset{\wedge}{y}_h$ so accurately that we can simply equate $\underset{\wedge}{y}_h$ with $\underset{\wedge}{x}_j$ by fixing the *h*th component function $y_h = \phi_{ah}(X,E)$ in (20) to be $y_h = 0\cdot x_1 + \ldots + 0\cdot x_{j-1} + 1\cdot x_j + 0\cdot x_{j+1} + \ldots + 0\cdot x_m + 0\cdot e_1 + \ldots + 0\cdot e_n$ or its equivalent. (This is what was meant above by describing certain Φ_a-components in our first version of the basic General Linear Model as "trivial.") But at the other extreme, we may conjecture virtually nothing about $\underset{\wedge}{X}$ beyond that some such *m*-tuple of hidden sources seems needed to explain the patterning of sample data on $\underset{\wedge}{Y}$ in *D*.

With suitable specification of its open parameters, law-conjecture (20) proffers explanation for the $\underset{\wedge}{Y}$-scores of single objects in *D*. but to analyze/interpret sample data on $\underset{\wedge}{Y}$, we must expand micro-model (20) into a macro-model (22) under which an aggregate of $[\underset{\wedge}{X},\underset{\wedge}{E}]$-events explains the $\underset{\wedge}{Y}$-state of a *group* of *D*-members. There is nothing at all mysterious about this expansion; but its technicalities even for random sampling are somewhat more intricate than commonsense kens, and getting clear on how SLese manages this is an essential preface to understanding the structure of more complex micro-molar systems. Given *N* individuals sampled from *D*, our first move is to index these from 1 to *N* in arbitrary but thereafter fixed order, and then to think of this collection as a molar unit, s_N, comprising *N* disjoint parts that can be picked out of s_N by indexical descriptors. That is, 'the *k*th individual in sample s_N', abbreviated $f_k(s_N)$ in (21), refers to sample-member o_k. This allows our observed array $\{\ulcorner y_j;o_k\urcorner: j = 1,\ldots,m;\ k = 1,\ldots,N\}$ of scattered micro-events to be formalized as a single macro-event $\ulcorner\underset{\wedge}{Y}^*;s_N\urcorner$ whose manifest locus is molar object s_N. Specifically, we multicopy each component $\underset{\wedge}{y}_j$ of $\underset{\wedge}{Y}$ into *N* t-derivative micro-variables $\underset{\wedge}{y}^*_{jk} =_{\text{def}} [\underset{\wedge}{y}_j f_k]$ ($k = 1,\ldots,N$) over the domain D^*_N of *N*-member samples from *D*—i.e., the value of $\underset{\wedge}{y}^*_{jk}$ for s_N is the value of $\underset{\wedge}{y}_j$ for s_N's *k*th member—and then compound all the latter as $\underset{\wedge}{Y}^* =_{\text{def}} [\underset{\wedge}{y}^*_{jk}: j = 1,\ldots,n;\ k = 1,\ldots,N]$. We can then conjecture that observed macro-event $\ulcorner\underset{\wedge}{Y}^*;s_N\urcorner$ is accounted for by a macro-antecedent $\ulcorner[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*];\underline{s}_N\urcorner$ under a macro-principle (22) derived from the micro-principle (20) under which, for each part $\underline{o}_k$ of $\underline{s}_N$, $\ulcorner\underset{\wedge}{Y};\underline{o}_k\urcorner$ is due to a micro-event $\ulcorner[\underset{\wedge}{X},\underset{\wedge}{E}];\underline{o}_k\urcorner$ that differs only by translocation from a fragment of $\ulcorner[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*];\underline{s}_N\urcorner$. Specifically, (22)'s transducer is a doubly indexed array $\Phi^*_a =_{\text{def}} \{\phi^*_{jk}: \underline{j} = 1,\ldots,\underline{n};\ \underline{k} = 1,\ldots,\underline{N}\}$ wherein each ϕ^*_{jk} is a function on the range of $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$ that outputs the $\underline{k}$th sample-member's value of the $\underline{j}$th data dimension by first selecting, from the macro-compound $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*](\underline{s}_N)$ of micro-input scores for all the sample members, just the sub-array contributed by $\underline{o}_k$ and then applying to this selection the $\underline{j}$th component ϕ_{aj} of transducer Φ_a in (20). If (20) is causal, (22) too is causal--or so the metaprinciple of Output Compounding claims--so long as the determination of each $\ulcorner\underset{\wedge}{y}_j;\underline{o}_k\urcorner$ by $\ulcorner[\underset{\wedge}{X},\underset{\wedge}{E}];\underline{o}_k\urcorner$ is unmediated by any micro-input event $\ulcorner\underset{\wedge}{x}_h;\underline{o}_{k'}\urcorner$ or $\ulcorner\underset{\wedge}{e}_j;\underline{o}_{k'}\urcorner$ whose locus is a member $\underline{o}_{k'}$ of $\underline{s}_N$ different from $\underline{o}_k$. (This mediation premise is reasonable enough when $\underline{s}_N$ intuitively qualifies as an "independent random sample" from $\underline{D}$; but it is generally less plausible for more natural molar systems discussed later.)

Elaboration of (20) into (22) is not an end in itself. Indeed, this would have little point were it not for the molar abstractions to which (22) gives rise. For it is now possible that we can a-derive from $\underset{\wedge}{Y}^*$ a compound molar pattern variable $\underset{\wedge}{\tilde{Y}}^*$ whose determination by ~~an~~ abstraction $\underset{\wedge}{\tilde{Z}}^*$ from $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$ under a transducer Γ^*_a entailed by Φ^*_a is sufficiently simple, given that the residuals compounded in $\underset{\wedge}{E}^*$ behave as statistical theory deems proper, that parameters $\underline{a}$ in Φ_a, Φ^*_a, and Γ_a can be recovered from the value of $\underset{\wedge}{\tilde{Y}}^*$ computed for $\underline{s}_N$ from data array $\underset{\wedge}{Y}^*(\underline{s}_N)$.

¶Specifically, let $\underline{G} = \langle \underline{g}_1,\ldots,\underline{g}_r\rangle$ be some array of numerical statistics (i.e. pattern abstractors) on experimental set-up $\underset{\wedge}{Y}^*$. (See p. 27f. for review of this terminology.) That is, each $\underline{g}_j$ is some number-valued function on the range of $\underset{\wedge}{Y}^*$, so that for each value $\underline{Y}^*$ of $\underset{\wedge}{Y}^*$, $\underline{g}_j\underline{Y}^*$ is a single number distilled by $\underline{g}_j$ out of score array $\underline{Y}^*$. Then from (22) it follows by Output Abstraction that

$$(23)\qquad \text{In } \underline{D}^*_N,\quad \underline{g}_j\underline{Y}^* = \underline{g}_j\,\Phi^*_a(\underline{X}^*,\underline{E}^*)\qquad (\underline{j} = 1,\ldots,\underline{r})\ ,$$

wherein g_jY^* is the value of pattern variable $[g_j\underset{\wedge}{Y}^*]$ for an arbitrary D_N^*-member whose value of $\underset{\wedge}{Y}^*$ is Y^*. If the component functions in tuples G and Φ_a are algebraically well-behaved in the fashion typified by polynomials, there will exist a finite tuple $H = \langle h_1,\ldots,h_s\rangle$ of statistics on $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$, identified just by G and the known form $\underset{\sim}{\Phi}$ of Φ_a, and for each g_j in G a function γ_{ja} on the range of H selected from a known function-form $\underset{\sim}{\gamma}_j$ by the same parameters $a = \langle\ldots,a_h,\ldots\rangle$ that select Φ_a from form $\underset{\sim}{\Phi}$, such that

$$g_j\Phi_a^* \;=\; \gamma_{ja}H \qquad (\,j = 1,\ldots,r\,)\ .$$

This rewriting of $g_j\Phi_a^*$ as $\gamma_{ja}H$ is nothing more than an algebraic rearrangement of elementary mathematical operations which can be combined in alternative ways to achieve the same mapping from the range of $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$ into the range of statistic g_j.

⟦Roughly speaking, the reorganization of $g_j\Phi_a^*$ sought by $\gamma_{ja}H$ is for H to perform as much condensation of the function's argument as is possible (under the constraint that H is to be the same for all $g_1,\ldots,g_r$) without involving any of the unknown parameters a, before γ_{ja} carries out the final distillation to which a is essential. To illustrate, let g_μ be the function that extracts the arithmetic mean of any N numbers to which it is applied, i.e., $g_\mu(z_1,\ldots,z_N) =_{\text{def}} (z_1 + \ldots + z_N)/N$. And write $\mathcal{L}_a^*$ for the linear transducer in (22.1). Then $\mathcal{L}_a^*$ is composable into g_μ (since all values of $\mathcal{L}_a^*$ are number N-tuples); and to see how their composition can be reorganized in the fashion desired, we apply $g_\mu\mathcal{L}_a^*$ to an arbitrary argument $\langle X^*,e^*\rangle$ and observe from the righthand side of (22.1) that

$$\begin{aligned} g_\mu\mathcal{L}_a^*(X^*,e^*) &= g_\mu(a_0 1^* + a_1x_1^* + \ldots + a_mx_m^* + e^*) \\ &= a_0 + a_1g_\mu(x_1^*) + \ldots + a_mg_\mu(x_m^*) + g_\mu(e^*)\ , \end{aligned}$$

where each $g_\mu(x_j^*)$ is the mean of the jth row x_j^* of score matrix X^*. If γ_a is now the linear compositor on number $(m+1)$-tuples whose value for any argument $\langle z_1,\ldots,z_{m+1}\rangle$ is $\gamma_a(z_1,\ldots,z_{m+1}) =_{\text{def}} a_0 + a_1z_1 + \ldots + a_mz_m + z_{m+1}$, this becomes simply

$$
\begin{aligned}
g_\mu f^*_a(\underline{X}^*,\underline{E}^*) &= \gamma_a(g_\mu(\underline{x}^*_1),\ldots,g_\mu(\underline{x}^*_m),g_\mu(\underline{e}^*)) \\
&= \gamma_a \underline{H}_\mu(\underline{X}^*,\underline{e}^*) ,
\end{aligned}
$$

where $\underline{H}_\mu$ is the compound function $\underline{H}_\mu = \langle g_\mu\sigma_1,\ldots,g_\mu\sigma_m,g_\mu\sigma_{m+1}\rangle$ each component $g_\mu\sigma_j$ of which is the composition into g_μ of a sub-array selector σ_j that picks out of score array $\langle\underline{X}^*,\underline{e}^*\rangle$ either $\underline{e}^*$ (when $\underline{j} = \underline{m}+1$) or the $\underline{j}$th row of $\underline{X}^*$. To appreciate how the composition of f^*_a into g_μ differs from that of $\underline{H}_\mu$ into γ_a even though $g_\mu f^*_a = \gamma_a\underline{H}_\mu$, note that $\underline{H}_\mu$ condenses the $(\underline{m}+1)\times\underline{N}$ element array $\langle\underline{X}^*,\underline{E}^*\rangle$ into just $\underline{m}+1$ numbers unaffected by parameters $\underline{a}$ whose force is then concentrated into how γ_a composites this reduced array, whereas condensation of $\langle\underline{X}^*,\underline{e}^*\rangle$ by the sequence of operations in $g_\mu f^*_a$ essentially exhausts the force of parameters $\underline{a}$ before any reduction in array size is accomplished.]

Given equivalences $\{g_j f^*_a = \gamma_{ja}\underline{H}\}$, (23) is mathematically equivalent to

(23$\underline{a}$) In $\underline{D}^*_N$, $\{g_j\underline{Y}^* = \gamma_{ja}\underline{H}(\underline{X}^*,\underline{E}^*)\}$ $(\underline{j} = 1,\ldots,\underline{r})$,

which by Output Compounding is

(24) In $\underline{D}^*_N$, $\underline{G}(\underline{Y}^*) = \Gamma_a(\underline{H}(\underline{X}^*,\underline{E}^*))$ $(\Gamma_a =_{\text{def}} \langle\gamma_{1a},\ldots,\gamma_{ra}\rangle)$,

where for any tuple $\underline{G} = \langle g_1,\ldots,g_r\rangle$ of functions on the range of a function or variable $\mathfrak{z}$, we continue to write $\underline{G}(\mathfrak{z}) =_{\text{def}} \langle g_1\mathfrak{z},\ldots,g_r\mathfrak{z}\rangle$. To make clear what (24) accomplishes, let $\underset{\wedge}{\tilde{Y}} = [\underset{\wedge}{\tilde{y}}_1,\ldots,\underset{\wedge}{\tilde{y}}_r]$ and $\underset{\wedge}{\tilde{Z}} = [\underset{\wedge}{\tilde{z}}_1,\ldots,\underset{\wedge}{\tilde{z}}_s]$ be the compound variables over $\underline{D}^*_N$ respectively a-derived from $\underset{\wedge}{Y}^*$ and $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$ by functions $\underline{G} = \langle g_1,\ldots,g_r\rangle$ and $\underline{H} = \langle\underline{h}_1,\ldots,\underline{h}_s\rangle$ in (24). That is,

$$\underset{\wedge}{\tilde{y}}_j =_{\text{def}} [g_j\underset{\wedge}{Y}^*] \quad (\underline{j} = 1,\ldots,\underline{r}), \qquad \underset{\wedge}{\tilde{z}}_k =_{\text{def}} [\underline{h}_k[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]] \quad (\underline{k} = 1,\ldots,\underline{s}) .$$

Then by Input Abstraction, (24) becomes simply

(24$\underline{a}$) In $\underline{D}^*_N$, $\tilde{\underline{Y}} = \Gamma_a(\underline{Z})$,

wherein molar transducer Γ_a is selected by the same parameters $\underline{a}$ as before from a known form $\underset{\sim}{\Gamma}$, $\underset{\wedge}{\tilde{Y}}$ is a compound pattern variable whose value for sample $\underline{s}_N$ is computable from data distribution $\underset{\wedge}{Y}^*(\underline{s}_N)$, and the value of compound molar variable $\underset{\wedge}{Z}^*$ for $\underline{s}_N$ is an abstraction from the unknown distribution of scores on $\underset{\wedge}{X}$ and $\underset{\wedge}{E}$ in $\underline{s}_N$. Similarly but not quite identically, equation (24) says that the value of compound statistic $\underline{G}$ for the observed $\underset{\wedge}{Y}$-distribution in $\underline{s}_N$ derives from--indeed, according to the metaprinciple of Input Abstraction, may in some sense be caused by--the value of compound statistic $\underline{H}$ for the sample distribution on $[\underset{\wedge}{X},\underset{\wedge}{E}]$ under a transducer identifiable by the same parameters $\underline{a}$ that identify transducer Φ_a in model (20) for production of $\underset{\wedge}{Y}$-scores for single individuals in $\underline{D}$.

A very simple example of (24$\underline{a}$) that follows from micro-model (20.1) by applying univariate statistic $\underline{g}_\mu$ to both sides of (22.1) is

$$(24.1) \qquad \text{In } \underline{D}^*_N, \quad \underline{\bar{Y}} = \underline{a}_0 + \underline{a}_1\underline{\bar{x}}_1 + \ldots + \underline{a}_m\underline{\bar{x}}_m + \underline{\bar{e}} \qquad (\text{ given } (20.1)\ ),$$

the overscores therein being standard notation for sample means on the variables indicated. But in the most prevalent applications of (24/24$\underline{a}$), abstractors $\underline{G}$ and $\underline{H}$ condense sample distributions into within-sample covariances. These applications begin with multiple-output linear micro-model

$$(20.2) \qquad \text{In } \underline{D}, \quad \underline{Y} = \underset{\sim}{\underline{A}}\underline{X} + \underline{E}\ ,$$

wherein $\underline{Y}$, $\underline{X}$, and $\underline{E}$ are column vectors of scores for an arbitrary individual in $\underline{D}$ on variables $\underset{\wedge}{Y} = [\underset{\wedge}{y}_1,\ldots,\underset{\wedge}{y}_n]$, $\underset{\wedge}{X} = [\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m]$, and $\underset{\wedge}{E} = [\underset{\wedge}{e}_1,\ldots,\underset{\wedge}{e}_n]$, respectively. (For simplicity, we suppress explicit recognition of this model's additive parameters. Those can be made zero by scaling conventions, or put into $\underset{\sim}{A}$ by including in $\underset{\wedge}{X}$ a pseudo-variable on which every $\underline{D}$-member has a score of unity.) The transducer in (20.2) is a linear vector function identified by a coefficient matrix $\underset{\sim}{A}$, the unknown elements of which instantiate generic parameter tuple $\underline{a}$ in (20). By the generic derivation of (22) from (20), it follows from (20.2) that the $\underset{\wedge}{Y}$-score distribution in any size-$\underline{N}$ sample $\underline{s}_N$ of $\underline{D}$-members has causal determination

(22.2) In $\underline{D}^*_N$, $\underline{Y}^* = \underset{\sim}{A}\underline{X}^* + \underline{E}^*$ (given (20.2))

wherein the generic $\{\sigma_k\}$-demarked difference between micro-transducer Φ_a in (20) and its macro-counterpart Φ^*_a in (22) is concealed within (22.2) by the rules of matrix multiplication. By applying Covariance abstractor $\underline{Cov}$ (see p. 28 above) to (22.2) we derive one of the classically fundamental theorems of linear data analysis, namely,

(24.2) In $\underline{D}^*_N$, $\underline{C}_{YY} = \underset{\sim}{A}\underline{C}_{XX}\underset{\sim}{A}^T + \underset{\sim}{A}\underline{C}_{XE} + (\underset{\sim}{A}\underline{C}_{XE})^T + \underline{C}_{EE}$ (given (20.2)) ,

wherein super-T denotes matrix-transpose. Each symbol of form $\underline{C}_{ZZ'}$ in (24.2) ($\underset{\wedge}{Z}$ and $\underset{\wedge}{Z}'$ variously being $\underset{\wedge}{Y}$, $\underset{\wedge}{X}$, or $\underset{\wedge}{E}$) stands for a numerical matrix, derived from the matrices $\underline{Z}^*$ and $\underline{Z}'^*$ of sample scores on variables $\underset{\wedge}{Z}$ and $\underset{\wedge}{Z}'$ in an arbitrary $\underline{s}_N$, whose $\underline{jk}$th element is the covariance between the $\underline{j}$th row of $\underset{\wedge}{Z}^*(\underline{s}_N)$ and the $\underline{k}$th row of $\underset{\wedge}{Z}'^*(\underline{s}_N)$. It would be tedious and probably unedifying to set out in fine detail how (24.2) instantiates generic formula (24/24$\underline{a}$). The salient point is simply that each matrix $\underline{C}_{YY}$, $\underline{C}_{XX}$, $\underline{C}_{XE}$, and $\underline{C}_{EE}$ in this enormously powerful equation is just the value for an arbitrary sample $\underline{s}_N$ of a compound pattern variable $\underset{\wedge}{C}_{YY}$, etc., defined by applying certain abstractor functions to sample-distribution variables $\underset{\wedge}{Y}^*$ and $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*]$. Morever, it is natural to claim here that output covariances $\underline{C}_{YY}$ are in some molar sense $\underline{\text{due}}$ $\underline{\text{to}}$ input covariances $\underline{C}_{XX}$, $\underline{C}_{XE}$, and $\underline{C}_{EE}$ under transduction (24.2), even though $\underline{C}_{YY}$ is an analytic abstraction from the score array $\underset{\wedge}{Y}^*(\underline{s}_N)$ which is caused by input distribution $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*](\underline{s}_N)$ without any literal causal mediation by the input covariances in this determination sequence.

Insomuch as molar model (24$\underline{a}$) has the same mathematical character as micro-model (20), we still need to consider why, for recovery of the unknown parameters in both (20) and (24$\underline{a}$), one is preferable to the other. Prima facie the reason seems evident: For good-sized $\underline{N}$, there is vastly more information available in macro-datum $\ulcorner\underset{\wedge}{Y}^*;\underline{s}_N\urcorner$ or its molar summary $\ulcorner\underset{\wedge}{\check{Y}};\underline{s}_N\urcorner$ (i.e. $\ulcorner\underline{G}(\underset{\wedge}{Y}^*);\underline{s}_N\urcorner$) for astutely chosen $\underline{G}$ than in any one micro-datum $\ulcorner\underset{\wedge}{Y};\underline{o}\urcorner$; and the $\underline{G}$-abstractors can discern important pattern features in $\ulcorner \underset{\wedge}{Y}^*;\underline{s}_N \urcorner$, notably correlations, that have no counterparts at all among the properties of any one micro-event $\ulcorner \underset{\wedge}{Y};\underline{o} \urcorner$. But (24$\underline{a}$)'s informational enrichment on the left is counterbalanced by proliferation of unknowns on the right; so a special magic is needed to make (24$\underline{a}$) more solvable than (20).

Sampling theory achieves this feat by interpreting orthodox sampling constraints on $\underline{s}_N$ to imply that almost all input components in (24$\underline{a}$) that abstract all or in part from residual distribution $\underset{\wedge}{E}^*(\underline{s}_N)$ are decently approximated by known constants or meagerly parameterized functions of $\underset{\wedge}{\tilde{Z}}(\underline{s}_N)$'s remainder $\underset{\wedge}{\tilde{Z}}_1(\underline{s}_N)$, so that for a small, perhaps zero, number of unknown parameters $\underline{b}$ in some known function-form $\underset{\sim}{\theta}$, $\underline{\tilde{Z}} \simeq \theta_b(\underline{\tilde{Z}}_1)$ in $\underline{D}^*_N$. [For more advanced models further constrained by (22-add), molar abstractions from the latter also contribute substantially to reduction $\underline{\tilde{Z}} \simeq \theta_b(\underline{\tilde{Z}}_1)$ of the number of input unknowns in (24$\underline{a}$).] Substituting this thinning of $\underset{\wedge}{\tilde{Z}}$ into (24$\underline{a}$) then approximately determines $\underset{\wedge}{\tilde{Y}}$ in $\underline{D}^*_N$ by a much smaller number of unknowns, namely,

$$(25) \qquad \text{In } \underline{D}^*_N, \quad \underline{\tilde{Y}} \simeq \Gamma_a(\theta_b(\underline{\tilde{Z}}_1)) \qquad (\underset{\wedge}{\tilde{Z}}_1 \text{ a subtuple of } \underset{\wedge}{\tilde{Z}}) .$$

And sampling theory also assures us that these approximations become increasingly exact, in approximately known degree, as sample size $\underline{N}$ becomes increasingly large.

> <u>Example</u>. When $\underline{G}$ and $\underline{H}$ abstract within-sample covariances for applications of the General Linear Model, standard sampling assumptions about E imply that all covariances in $\underline{C}_{XE}$ and all off-diagonal ones in $\underline{C}_{EE}$ converge with increasing $\underline{N}$ to zero, thereby simplifying (24.2) to
>
> $$(25.2) \qquad \text{In } \underline{D}^*_N, \quad \underline{C}_{YY} \simeq \underset{\sim}{A}\underline{C}_{XX}\underset{\sim}{A}^T + \underline{D}_E ,$$
>
> wherein $\underline{D}_E$ is the diagonal matrix of $\underset{\wedge}{E}$-component variances.

Not only is the ratio of unknowns-on-the-right to knowns-on-the-left for (25) generally lower than for any of (24$\underline{a}$), (22), or (20), it may well be low enough (if parameterization of (20) has been sufficiently frugal) to permit solution of (25) for all its righthand terms.

That is, under favorable mathematical circumstances which data-analytic practice labors mightily to arrange by its sampling assumptions, choice of transducer form Φ in micro-model (20), and selection of the experiment's summary statistics $\underline{G}$, the known form of (25)'s transducer, viewed as a function whose argument is all of $\langle \underline{a},\underline{b},\underline{\tilde{Z}}_1\rangle$, will have a computable inverse. Transforming both sides of (25) by this inverse, or by some approximation thereto designed to minimize the error resulting from the imperfect equality in (25), then effectively estimates both $\langle \underline{a},\underline{b}\rangle$ and the $\underline{H}$-abstracted pattern properties of unobserved macro-event $\ulcorner[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*];\underline{s}_N\urcorner$ from the $\underline{G}$-patterning in observed macro-event $\ulcorner\underset{\wedge}{Y}^*;\underline{s}_N\urcorner$. Moreover, if $\underline{N}$ is reasonably large, we will have reason for confidence that the true values of these statistics are close to our so-computed estimates thereof.

⟦In practice, it often occurs that whatever specialization of (20) seems most substantively attractive for particular variables $\underset{\wedge}{Y}$ is "underdetermined" in that its molar derivative (25) has no unique solution for $\langle \underline{a},\underline{b}\rangle$ under any choice of statistics $\langle \underline{G},\underline{H}\rangle$. This problem can always be solved superficially by placing additional constraints on $\langle \underline{a},\underline{b}\rangle$; but that in effect identifies only a class of admissible solutions for $\langle \underline{a},\underline{b}\rangle$ corresponding to our alternatives for those constraints. The deeper challenge of underdetermined models--importantly illustrated by the factor-analytic literature on terminal positioning of factor axes--is to find grounds on which to consider some of these admissible solutions more plausible than others as estimates of parameters in causal transducers by which Y-data are produced. Note also that solution for all unknowns in (25) identifies the "structural" parameters $\underline{a}$ in model (20/22/24$\underline{a}$) without recovering the underlying input distribution on $[\underset{\wedge}{X},\underset{\wedge}{E}]$ in sample $\underline{s}_N$. Even when $\Phi_{\underline{a}}$ is fully known, inference of $[\underset{\wedge}{X}^*,\underset{\wedge}{E}^*](\underline{s}_N)$ from $\underset{\wedge}{Y}^*(\underline{s}_N)$ has its own special, only semi-tractable, difficulties that have likewise generated a considerable multivariate literature under the title "factor indeterminacy."⟧

Sampling theory's parting beneficence for applied data analysis/interpretation is to set out with some precision the degrees of credence warranted by (25)-based estimates of a causal model's parameters given our model's background presumptions. Its most sophisticated ideal is to furnish Bayesian posterior credibilities for $\langle \underline{a},\underline{b} \rangle$ and $\underline{H}([\underset{\wedge}{\underline{X}}^*,\underset{\wedge}{\underline{E}}^*](\underline{s}_N))$; but in practice we are more likely to be offered, in order of increasing popularity and decreasing epistemic merit, likelihood functions, confidence intervals, or null-hypothesis tests. What sampling theory cannot yet cogently appraise, however, is the plausibility of our choice of model form in (20) or our warrant for the flagrantly hypothetical sampling presumptions on which our model solution so severely depends. The latter issues are still on the wilderness side of science's epistemological frontier.

The micro-origins of macro-phenomena.

A perennial controversy in the philosophy of science has been the extent to which the way things are in the large is constituted by their character in the small. Although we have established no grounds here for strong reductionistic theses, we do have well in hand the main conceptual resources needed to make these scientifically articulate. Specifically, definition of scientific variables by compounding and a/t-derivation, together with derivation of laws by the assorted metaprinciples reviewed earlier, especially Input Abstraction and Output Abstraction, gives us the machinery to assemble basic events and the primary principles that govern them into molar description and explanation of macro-systems. And the practical payoff of sampling statistics demonstrates how important this construction can be for technical science even though sampling theory is just the most transparently simple variety of molar explanation. "Reduction" of macro-phenomena to their micro-constituents is merely operation of this derivational machinery in reverse.

I have spoken freely here about "translocation" functions without saying much about their nature. Although any relation in which genuine causal loci participate can be used to define translocators of greater or lesser artificiality, those locus connections that occur most naturally in the antecedents of basic causal laws, and must hence be worked into translocators if manifest law-form (8) is to prevail over (9'), are (*a*) excursors, i.e. whatever displacements from the locus of a causal event select where-and-when that cause gives rise to effects, and (*b*) part/whole couplings. I shall say nothing here about the ontology of excursive displacements, nor much about that of part/whole relations either. But since we have reason to suspect that most, perhaps all, properties of commonsense objects derive from the properties of their parts, it is important to understand the logic by which such derivations are possible.

Let us say that an entity *o* is a *compound object of kind C* iff, for some integer *n*, *o* is an *n*-tuple and *C* is an *n*-ary relation (in general, one that is logically complex) that is exemplified by *o*.[15] More generally, we can let *o* be

---

[15] By speaking of "logically complex relations" here we expand our earlier working presumption that basic predicates signify properties (see p. 22f.) into treatment of any simple or complex predicate (i.e., open sentence) as signifying some abstract entity common to just those things that satisfy this predicate. Thus from the sentence 'John likes cheesburgers, Mary is blonde, and John is older than Mary', nominalizing the predicate that results from substitution of placeholders for 'John' and 'Mary' therein gives us the concept of an abstract entity, The-relation-in-which-something-that-likes-cheesburgers-is-older-than-something-that-is-blonde. That such abstract entities really exist whenever English grammar generates names for them merits considerable scepticism. But it is virtually impossible to run the business of technical science, mathematics, and much of philosophy without freely nominalizing complex predicates, so we can only hope that some future philosophical enlightenment will find a way to view these as ontologically innocuous. Neither have we any practical way to avoid treating tuples and other indexed sets as real objects distinct from their individual components. But tuples, too, are grounds for ontological concern. For if

$o = \langle o_1,\ldots,o_n\rangle$ is an n-tuple of distinct objects and $o' = \langle o'_1,\ldots,o'_n\rangle$ is is a proper permutation of o, then the very point of conceiving o and o' as "ordered" is for o' not to be identical with o. Yet at the same time, it is hard to argue convincingly that the reality designated e.g. by 'First John, then Mary' differs in any language-independent way from the referent of 'First Mary, then John' or for that matter from that of 'First John, then Mary, then John again'. One can only hope that ontological differences created by indexing will someday soften into harmlessly convenient figures of speech or, more felicitously perhaps, into relations between words and things rather than just things alone. Meanwhile, I shall say that two indexed sets containing the same elements under different indexing or more generally having the same mereological sum of parts, are "de re equivalent."

---

any indexed collection of parts; but indexing as a finite tuple is most convenient for characterizing o's object kind. (Whether finite indexing always suffices need not concern us ~~here.) We do not require a compound object's components to be~~ disjoint. Each component o_k of a kind-C compound object $o = \langle o_1,\ldots,o_n\rangle$ is itself generally a compound object of some kind C'_k that may or may not be the same for all $k = 1,\ldots,n$. In the special statistical case already discussed, any size-N statistical sample is a compound object of kind D^*_N for some condition D such that each member of the sample is of kind D; beyond that, there is little more to a sample's being of kind D^*_N than just comprising N distinct kind-D parts whose respective values of certain residual input variables have no common sources. (This vaguely stated parts-disconnection constraint does not fully capture the notion of "random sampling," but it does point to the main force of requiring a sample's members to be drawn from D "independently" of one another.) In contrast, "real" compound objects that commonsense views as the loci of macrophenomena--plants and animals, machines of all complexities from levers and pulleys to airplanes and computers, social groups from insect colonies through families and tribes to nations, spaceship Earth, even galaxies and beyond--consist of parts diversified in kind but strongly interactive as causal loci.

Specifically, for any natural compound object whose parts-parsing is sufficiently detailed, there generally exist many causal laws of form (9/9') (p. 36 above) whose domain preconditions $\mathcal{P}$ are satisfied by one or more subarrays of o's parts. When $\mathcal{P}$ includes an appreciable excursive displacement and o is just one stage s-at-t of an enduring subject s, it may also well occur that parts of *o* are γ-related to one or more parts of some successor $o' =$ *s*-at-$t+\Delta$ of *o*. Indeed, such excursive links between parts of *o* and parts of *o*'s successors are mandatory if *s*'s stages are to be integrated by causal recursions. But the space-time region occupied by compound object *o* = *s*-at-*t* is generally no more instantaneous in time than it is pointlike in space; and although our conception of *o*'s temporal boundaries is in most instances unconsionably vague, its excursive width is generally ample for many micro-events whose loci are parts of *o* to have micro-effects whose loci are also parts of *o* rather than of some *o*-successor fully separate from *o*.[16] And of course the downstream effects of events whose loci are parts of *o* can generally be translocated, when formally useful to do so, into *t*-derivative events manifestly located within *o*. But nevermind technicalities just yet; my prefatory point is simply that the multifarious parts of a natural macro-object *o* are variously interrelated in ways that satisfy the preconditions of micro-laws whose ensemble abstracts into the molar behavior of *o*'s kind. This internal *structure* of compound objects is a major theme in the SLese story to be told of macro-systems.

Structure aside, the logic by which compound objects of a particular kind derive their causal behaviors from that of their parts is largely evident in the sketch of this already given for statistical compounds. But we now need a generic way to identify specific selections of the object's parts in terms of their function within the whole relative to its kind. The hardest work for this identification

---

[16] Even when object *s*-at-*t* is treated formally as having zero temporal width, the molar properties we ascribe to it are often abstracted from an ensemble of micro-events spread in time around instant *t*. (Cf. *Obtaining-a-Wechsler-Bellevue-IQ-score-of-132*, *Solving-today's-crossword-puzzle*, *Ordering-a-pepperoni-pizza*, etc.) Although it is customary to think of an object's successive temporal stages as stacked like slices of a salami, a more perspicuous image is that of links in a chain. That is, the temporally aft parts of *s*-at-*t* are also generally fore-parts of *s*-at-$t+\Delta$ for sufficiently small Δ. (I regret relegating this point to a footnote; for if your thinking about causal systems implicitly presupposes the sliced-salami model of thing-stages, you have much to gain from meditation on its overlapping-segments alternative.)

is converting the unanalyzed whole into a collection of distinctively indexed parts, e.g., treating John as equivalent to some ordering of his heart, lungs, liver, kidneys, etc., or, at a finer grain, of the individual cells and connecting tissues therein, and from there finding a causally cogent kind-classification for the parsed object keyed to this indexing. I have already presupposed solutions to the initial parts-parsing problem by formalizing compound object $\underline{o}$ as an $\underline{n}$-tuple,[17] and its kind $\underline{C}$ as a complex $\underline{n}$-ary relation. From there, to pick out parts of kind-$\underline{C}$ objects in terms of their formal positions within the whole, we can simply take any tuple $\underset{\sim}{\mu}$ of positive integers to define a function $\mu(\_\_)$ that maps any $\underline{n}$-tuple $\underline{o} = \langle \underline{o}_1, \ldots, \underline{o}_n \rangle$ of object-parts into the tuple of $\underline{o}$'s parts indexed by $\underset{\sim}{\mu}$. Thus if $\underset{\sim}{\mu} = \langle 4,1,6 \rangle$ and John = <this heart, this left lung, this right lung, this liver, this left kidney, this right kidney, this stomach, ...>, μ(John) is the 3-tuple <this liver, this heart, this right kidney>. ⟦To cope technically with cases where $\underset{\sim}{\mu}$ contains integers larger than $\underline{n}$, we can treat each object $\underline{n}$-tuple $\underline{o}$ as continued by an infinite sequence of "null" parts that can formally be picked out of $\underline{o}$ by μ but which are then null extensions of the real string of $\underline{o}$'s parts in $\mu(\underline{o})$.⟧ Meanwhile, the functional role in kind-$\underline{C}$ objects formally identified by any such selector function μ lies in what $\underline{o}$'s being of kind $\underline{C}$ implies about the relational/nonrelational character of $\underline{o}$'s subarray $\mu(\underline{o})$.

Specifically, if the complex of properties/relations we take to define object-kind $\underline{C}$ is causally salient, i.e., if the assortment of conditions imposed by $\underline{C}(\_, \ldots, \_)$ on the various parts of its satisfiers $\underline{o} = \langle \underline{o}_1, \ldots, \underline{o}_n \rangle$ suitably discriminate the micro-causal behavior of those parts, then certain selector functions $\{\mu_k\}$ will pick out parts of $\underline{o}$ that satisfy the domain preconditions of one or more causal laws $\{\underline{L}_k\}$ consistently across all $\underline{C}$-kind objects. That is,

---

[17] Technically, this amounts to our having managed to define a mapping ζ of some set of pre-parsed molar objects $\{\omega\}$ into $\underline{n}$-tuples $\{\underline{o} = \zeta(\omega)\}$. Whether $\zeta(\omega)$ is an exhaustive compilation of ω's parts, or even whether the components of $\underline{o}$ are strictly parts of ω at all, matters only when, in the throes of reduction/emergence controversy we seek to equate properties of ω with those of $\zeta(\omega)$.

if we express these laws $\{L_k\}$ in canonical form ($8'$) by embedding their locus structures within the translocators of t-derivative variables, object-kind C will be characterized by principles of form

(26) For any compound object o, if o is of kind C then $\mu_k(o) \in D_k$,

in which D_k is the domain of some causal law

(26-L_k) In D_k, $Y_k = \phi_k(X_k)$.

(It will be evident how (26) generalizes the requirement on kind-D^*_N statistical samples that for each $k = 1,\ldots,N$, the kth member of each N-tuple in D^*_N is in domain D.) For any o of kind C, any subtuple $\mu_k(o)$ of o for which some principle of form (26) holds may be viewed as a "module" of o whose causal kind is D_k. And correspondingly, we may call any such function μ_k a "module selector." Module selection is a formally species of translocation even though its ontology is altogether different from that of excursive translocation.

[[Rather than stipulating that μ_k selects subtuple $\mu_k(o)$ of o by indexing, as adopted here, we could more generally allow a kind-C module selector to be any function μ_k defined by a description of form 'The tuple of ___'s parts satisfying conditions C_k', so long as some principle of form (26) is then true of μ_k. This broadening of module selection would make no difference for what is done with the concept here except for waiving the requirement that C-kind objects have been parsed as tuples. But conversely, any finite array of so-liberalized module selectors on a class C of holistically conceived macro-objects can be used to define a parsing of C's members as compound objects out of which these same modules can be picked by index selectors.]]

The quality of lawfulness in structural generalities of form (26) can be wonderously diverse. It can be all or in part logical necessity, by virtue of our including in C's definition some or all D_k-conditions as requirements on the μ_k-subtuples of kind-C objects. It may be partly causal in that some properties of $\mu_k(o)$ definitive of C cause $\mu_k(o)$ to have some or all of the properties that put $\mu_k(o)$ in D_k. And very likely it is grounded in part on infra-causal locus relations, notably, the geometry of space/time. We can even allow (26) to be an "accidental" generality, since that can always be converted into a logical truth simply by adding it to our definition of C without changing this kind's extension. Regardless of its status in the world's explanatory order, let us call any principle of form (26) an *assembly law*. The assembly laws for a given C tell how kind-C objects are put together functionally, i.e., how they are constructed out of parts that are variously disposed by character and connection to work together one way rather than another.

Of course, assembly laws (26) alone do not explain the behavior of kind-C objects; they only specify which portions thereof are modules of what causal kinds. To complete the story, we also need the conjugate causal laws (26-L_k) which govern modules of these kinds. ⟦Moreover, there are finer micro-structural details of these modules and their laws that explanation of C-kind molar behavior also needs to recognize, especially in view of our formalizing module functioning by schema (8) rather than by the more articulate (9/9'). To illustrate, suppose for each o in C that o's modules $\mu_j(o)$ and $\mu_k(o)$ are both in the domain D of a form-(8) law whose t-derivational analysis shows it to have embedded locus structure

$$L_y: \qquad \text{In } D,\quad yf_0 = \phi(yf_1, yf_2, e) \quad .$$

(For example, when each o in C is some society whose modules include breeding couples $\{b = \langle b_1, b_2\rangle\}$ that are among the 2-tuples in dyad domain D, $f_1(b)$ and $f_2(b)$ might be respectively the male and female component of breeding couple b, with $f_0(b)$ the couple's first clild and L_y telling how parents' y-values determine the y-value of their first child.) Then we want to know whether micro-object $f_0\mu_j(o)$ is identical with $f_1\mu_k(o)$ or with $f_2\mu_k(o)$ for each o in C. (E.g., it might be a condition on societies parsed as tuples of kind $\underline{C}$ that the first child of couple $\mu_j(\underline{o})$ is one of the parents in couple $\mu_k(\underline{o})$.) For if that is so, then there is an auto-regression on $\underline{L}_y$ within kind-$\underline{C}$ objects by virtue of which events $\ulcorner \underset{\wedge}{y};\underline{f}_1\mu_j(\underline{o})\urcorner$ and $\ulcorner \underset{\wedge}{y};\underline{f}_2\mu_j(\underline{o})\urcorner$ affect event $\ulcorner \underset{\wedge}{y};\underline{f}_0\mu_k(\underline{o})\urcorner$ through the mediation of event $\ulcorner \underset{\wedge}{y};\underline{f}_0\mu_j(\underline{o})\urcorner$. Such causal recursions among the micro-events underlying the molar properties of kind-$\underline{C}$ objects establish regularities in $\underline{C}$ that cannot be derived just from $\underline{C}$'s assembly laws and within-module functioning.]]

Putting together these varied facets of a compound objects' micro-structure gives us

Definition 2. A structural analysis of compound object-kind $\underline{C}$ is any 4-tuple $\langle \underset{\sim}{P},\underset{\sim}{L},\underset{\sim}{R},\underset{\sim}{k}\rangle$ in which: (1) $\underset{\sim}{k}$ is an index set. (2) $\underset{\sim}{P}$ and $\underset{\sim}{L}$ are $\underset{\sim}{k}$-indexed sets $\underset{\sim}{P} = \{\underline{P}_k\colon \underline{k} \in \underset{\sim}{k}\}$ and $\underset{\sim}{L} = \{\underline{L}_k\colon \underline{k} \in \underset{\sim}{k}\}$ of assembly laws and causal laws respectively in which, for each $\underline{k}$ in $\underset{\sim}{k}$ and some module selector μ_k, $\underline{P}_k$ and $\underline{L}_k$ are of respective forms

$\underline{P}_k$: For each $\underline{o}$ of kind $\underline{C}$, $\mu_k(\underline{o}) \in \underline{D}_k$,

$\underline{L}_k$: In $\underline{D}_k$, $\underline{Y}_k = \phi_k(\underline{X}_k)$.

[[(3) $\underset{\sim}{R}$ is a pair $\underset{\sim}{R} = \langle \underset{\sim}{R}_t,\underset{\sim}{R}_s\rangle$ wherein: (3_t) $\underset{\sim}{R}_t$ is a set

$$\left\{ \underset{\wedge}{z}_h = [\underset{\wedge}{z}'_h\underline{f}_h] \right\}$$

of t-derivational analyses of some or all of the variables in $\underset{\sim}{L}$, i.e., each $\underset{\wedge}{z}_h$ is one of the variables in some law $\underline{L}_k$ in $\underset{\sim}{L}$; and (3_s) $\underset{\sim}{R}_s$ is a set

$$\left\{ \text{For all } \underline{o} \text{ in } \underline{C},\quad \lambda_{hh'kk'}(\underline{f}_h\mu_k(\underline{o})) = \nu_{hh'kk'}(\underline{f}_{h'}(\mu_{k'}(\underline{o})) \right\}$$

of locus-identity principles wherein μ_k and $\mu_{k'}$ are module selectors used in $\underset{\sim}{P}$ to pick out subtuples from kind-$\underline{C}$ objects, $\underline{f}_h$ and $\underline{f}_{h'}$ are either Identity functions or translocators disclosed in t-derivational analysis $\underset{\sim}{R}_t$ of $\underset{\sim}{L}$, and

$\lambda_{hh'kk'}$ and $\gamma_{hh'kk'}$ are subtuple selectors that pick out (possibly null) identical parts of $f_h\mu_k(o)$ and $f_{h'}\mu_{k'}(o)$, respectively.⟧

For any o of kind C, we may say that structural analysis $\langle \mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ of C also describes the *micro-structure of o at level C* wherein $\mathbf{P}$ and $\mathbf{L}$ give o's C-level *assembly structure* and *micro-causal structure*, respectively, ⟦while $\mathbf{R}_s$ and $\mathbf{R}_t$ are *micro-identity supplements* to $\mathbf{P}$ and $\mathbf{L}$, respectively.⟧

Notes:

1) Structural analysis of object-kind C is nonunique; trivially because this can be indexed in different ways, and nontrivially because Def. 2 does not stipulate that a structural analysis must be complete. (Structural analysis $\langle \mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ of C is *complete* iff, for every other structural analysis $\langle \mathbf{P}',\mathbf{L}',\mathbf{R}',\mathbf{k}'\rangle$ of C and every index k' in $\mathbf{k}'$, there is an index k in $\mathbf{k}$ such that $P'_k = P_k$ and $L'_k = L_k$ while $\mathbf{R}'$ is similarly embedded in $\mathbf{R}$.) Even so, we may speak heuristically of *the* structural analysis of C with the understanding that this is complete with some uniquely specified indexing.

2) The assembly structure of a complete structural analysis of object-kind C is essentially the "causal nature" of C, while the assembly structure of any incomplete structural analysis of C is a partial causal nature of C. This is because if $\langle \mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ is a complete structural analysis of C, $\mathbf{P}$ supplemented by $\mathbf{R}_s$ identifies all that is relevant in C for the causal functioning of C-kind objects. Specifically, if P^* is the conjunction of all predicates $\{\mu_k(\_\_) \in D_k : k \in \mathbf{k}\}$ in P, each object of kind C is also of kind P^*; if $\langle \mathbf{P}',\mathbf{R}'\rangle$ differs from $\langle \mathbf{P},\mathbf{R}\rangle$ just by replacing all occurrences of C in the latter by P^* then $\langle \mathbf{P}',\mathbf{L},\mathbf{R}',\mathbf{k}\rangle$ is a structural analysis of P^*; and any object in P^* but not in C in fact shares with kind-C objects all micro-causal properties common to the latter.

3) Def. 2 presumes that all laws and identity-claims in a structural analysis $\langle \mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ of object-kind C are *true* (so that whenever ontologically

tolerable, $\langle\mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ can alternatively be viewed as comprising indexed sets of objective facts rather than assertions). Otherwise, if $\mathbf{P}$, $\mathbf{L}$, and $\mathbf{R}$ in Def. 2 contain merely *laws and *identities, $\langle\mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ is a structural *analysis of object-kind D that merely conjectures how kind-C objects *may* work causally.

4) The sense in which the structural analysis of object-kind C is also a structural analysis of any particular o of this kind is that from $\langle\mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ and the information that o is of kind C, it follows for each k in $\mathbf{k}$ that o's module $\mu_k(o)$ satisfies domain conditions D_k and hence, from L_k, is caused by event $\ulcorner \underset{\wedge}{X}_k;\mu_k(o)\urcorner$ to have value $\phi_k(\underset{\wedge}{X}_k(\mu_k(o)))$ of variable $\underset{\wedge}{y}_k$. Strictly speaking it is these latter facts specifically about o, rather than generalities $\mathbf{P}$ and $\mathbf{L}$, that constitute the level-C assembly structure and micro-causal structure of any particular kind-C object o.

5) When $\langle\mathbf{P},\mathbf{L},\mathbf{R},\mathbf{k}\rangle$ is the structural analysis of object-kind C, and j,k are both indices in $\mathbf{k}$, $j \neq k$ does not preclude $\mu_j = \mu_k$, $D_j = D_k$, or $L_j = L_k$. That is, any given level-C module of o can be of many different causal kinds D_k (albeit we may take *the* causal kind of $\mu_k(o)$ to be the conjunction of all of these), and different modules of o can be of the same kind. To make explicit the multiplicity of causal laws that may simultaneously govern o's module $\mu_k(o)$, L_k in Def. 2 can be generalized to read L_k: $Y_k = \Phi_k(X_k)$.

6) Structural analysis of macro-phenomena is highly relative to a discretionary "level" of analysis for two reasons. First of all, the parts-parsing of a holistically conceived molar object ω that prepares it for analysis under Def. 2 can be developed at many different grains from coarse to fine. For example, if ω is de re equivalent to $\langle o_1,o_2\rangle$, i.e. these have the same mereological sum of parts, while o_1 and o_2 are de re equivalent to $\langle o_{11},o_{12}\rangle$ and $\langle o_{21},o_{22}\rangle$, respectively, then ω is also de re equivalent to $\langle o_{11},o_{12},o_2\rangle$, to $\langle o_1,o_{21},o_{22}\rangle$, and to $\langle o_{11},o_{12},o_{21},o_{22}\rangle$. And secondly, even after arbitration of grain has identified ω with a particular n-tuple o of its parts, structural analysis of o is still relative to

a level $\underline{C}$ because kinds of $\underline{n}$-tuples for fixed $\underline{n}$ form a partial order in which some entail others. Thus if $\underline{C}$ and $\underline{C}'$ are different $\underline{n}$-tuple kinds, so is their union $\underline{C}$-or-$\underline{C}'$. The more determinate (i.e. restrictive) is the kind $\underline{C}$ to which we assign a given $\underline{o}$, the richer in general is the array of structural facts about $\underline{o}$ that follow from structural analysis of $\underline{C}$. Any particular compound object $\underline{o}$ is theoretically of a maximally determinate kind $\underline{C}^*$ such that for any other kind $\underline{C}$ to which $\underline{o}$ belongs, the complete structural analysis of $\underline{C}$ is a fragment of the complete structural analysis of $\underline{C}^*$ However, this $\underline{C}^*$ may well be unique to $\underline{o}$ and would hence be an inauspicious domain within which to seek broadly applicable generalities even if we had any practical way to recognize $\underline{C}^*$ in the first place.

Molar consequences of micro-structure.

It is an immediate consequence of Def. 2 that if $\langle \underset{\sim}{P}, \underset{\sim}{L}, \underset{\sim}{R}, \underset{\sim}{k} \rangle$ is a structural analysis of object-kind $\underline{C}$, then by Domain Translocation,

$$(27) \qquad \text{In } \underline{C},\ \left\{ \underline{Y}_{k}\mu_{k} = \phi_{k}(\underline{X}_{k}\mu_{k}) \right\} \qquad (\underline{k} \in \underset{\sim}{k})$$

is an ensemble of form-(8') causal laws whose domain is $\underline{C}$ for all $\underline{k}$ in $\underset{\sim}{k}$. ⟦And the micro-identities in $\underset{\sim}{R}$ generally reveal causal recursions among laws (27) that augment these with the products of integrable laws in (27) as well.⟧ So by Input Expansion and Output Compounding ⟦as well as Mediated Composition in light of $\underset{\sim}{R}$⟧, laws (27) generate an array of supervenient micro-laws of form

$$(28) \qquad \text{In } \underline{C},\ \underline{Y}_{j} = \Phi_{j}(\underline{Z}_{j})$$

in which each component of $\underset{\wedge}{Y}_j$ and $\underset{\wedge}{Z}_j$ differs only by translocation from some variable in $\underline{C}$'s micro-causal structure $\underset{\sim}{L}$.

Example. Suppose that two of the untranslocated micro-laws in $\underset{\sim}{L}$ are

$$\underline{L}_1\text{: In } \underline{D}_1,\ \underline{y}_1 = \phi_1(\underline{X}_1)\ ; \qquad \underline{L}_2\text{: In } \underline{D}_2,\ \underline{y}_2 = \phi_2(\underline{X}_2)\ ;$$

with the assembly structure $\underset{\sim}{P}$ of kind-$\underline{C}$ objects implying that μ_1 and μ_2 select modules of kind $\underline{D}_1$ and $\underline{D}_2$, respectively, from each $\underline{o}$ in $\underline{C}$. Then the fragment of (27) for $\underline{k} = 1,2$ is

$$\text{In } \underline{C},\ \left\{\begin{array}{lcl} \underline{y}_1\mu_1 & = & \phi_1(\underline{X}_1\mu_1) \\ \underline{y}_2\mu_2 & = & \phi_2(\underline{X}_2\mu_2) \end{array}\right\}\ ,$$

which by Input Expansion becomes

$$\text{In } \underline{C},\ \left\{\begin{array}{lcll} \underline{y}_1\mu_1 & = & \phi_1\sigma_1(\underline{X}_1\mu_1,\underline{X}_2\mu_2) & (\ \sigma_1(\underline{X}_1,\underline{X}_2) =_{\text{def}} \underline{X}_1\ ) \\ \underline{y}_2\mu_2 & = & \phi_2\sigma_2(\underline{X}_1\mu_1,\underline{X}_2\mu_2) & (\ \sigma_2(\underline{X}_1,\underline{X}_2) =_{\text{def}} \underline{X}_2\ ) \end{array}\right\}\ ,$$

and from there, by Output Compounding, gives the instantiation of (28) wherein $\underset{\wedge}{Y}_j = [\underset{\wedge}{y}_1\mu_1,\underset{\wedge}{y}_2\mu_2]$, $\underset{\wedge}{Z}_j = [\underset{\wedge}{X}_1\mu_1,\underset{\wedge}{X}_2\mu_2]$, and $\underline{\Phi}_j = \langle\phi_1\sigma_1,\phi_2\sigma_2\rangle$.

Any form-(28) compound law so derived from (27) is tantamount to some sub-collection of laws in (27)'s recursive unfolding; and index $\underline{j}$ in (28) tokens that many such selections are possible. Indeed, we can take $\underline{j}$ to range over all finite subtuples of the index set $\underset{\sim}{k}^+$ comprising $\underset{\sim}{k}$ augmented by indices for all products of integrable laws in (27), except that the law (28) output-compounded out of the input expansions of the laws picked out of the recursive unfolding of (27) by an arbitrary subtuple $\underline{j}$ of $\underset{m}{\underline{k}}^{+}$ may violate the mediation conditions under which laws derived by Input Expansion are causal. Derivation of (28) from (27) is illustrated by sampling theory's collection of equations (21) into single compound equation (22).

Compound law (28) entails molar phenomena in domain $\underline{C}$ by the same principles of abstraction that convert (28)'s instance (22) into statistical regularities (25). For any tuple $\underline{G}_j = \langle g_1, g_2, \ldots \rangle$ of abstractor functions on the range of compound variable $\underset{\wedge}{Y}_j$, let $\underline{H}_j = \langle \underline{h}_1, \underline{h}_2, \ldots \rangle$ be some tuple of abstractors on the range of $\underset{\wedge}{Z}_j$ such that $\underline{G}_j(\Phi_j) = \Psi_j(\underline{H}_j)$ for some tuple $\Psi_j = \langle \psi_{j1}, \psi_{j2}, \ldots \rangle$ of transducers on the range of $\underline{H}_j$. That is, let

$$g_i \Phi_j = \psi_{ji} \underline{H}_j \qquad (\underline{i} = 1,2,\ldots)$$

be some more-or-less arbitrary mathematical reorganization of transducers $\langle g_1\Phi_j, g_2\Phi_j, \ldots \rangle$. Then by Input Abstraction it follows from (28) that

(29) In $\underline{C}$, $\underline{G}_j(\underline{Y}_j) = \Psi_j(\underline{H}_j(\underline{Z}_j))$,

or equivalently

(29$\underline{a}$) In $\underline{C}$, $\underline{\tilde{Y}}_j = \Psi_j(\underline{\tilde{Z}}_j)$ $\left(\underset{\wedge}{\tilde{Y}}_j =_{\text{def}} \underline{G}_j(\underset{\wedge}{Y}_j),\ \underset{\wedge}{\tilde{Z}}_j =_{\text{def}} \underline{H}_j(\underset{\wedge}{Z}_j) \right)$,

is a tuple of laws in which molar events $\{\ulcorner \underset{\wedge}{\tilde{Z}}_j ; \underline{o} \urcorner\}$ $(\underline{o} \in \underline{C})$ are prima facie responsible, a some level of molar causality, for molar events $\{\ulcorner \underset{\wedge}{\tilde{Y}}_j ; \underline{o} \urcorner\}$ $(\underline{o} \in \underline{C})$. The tilde in $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{Z}}_j$ tokens that each component variable $\underset{\wedge}{\tilde{y}}_{ij} = [g_i \underset{\wedge}{Y}_j]$ in $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{z}}_{ij} = [\underline{h}_i \underset{\wedge}{Z}_j]$ in $\tilde{Z}_j$ can be viewed as some composite of the component variables in $\underset{\wedge}{Y}_j$ and $\underset{\wedge}{Z}_j$, respectively, albeit not necessarily one that is at all an "averaging" of its constituents. Laws of form (29$\underline{a}$) characterize the <u>macro-causal</u> behavior of object-kind $\underline{C}$, while the common dependency of output components in $\underset{\wedge}{\tilde{Y}}_j$ upon the same inputs $\underset{\wedge}{\tilde{Z}}_j$ induces a distributional patterning of scores on the various $\underset{\wedge}{\tilde{Y}}_j$-dimensions that can be discerned in standard statistical fashion within sufficiently large samples of kind-$\underline{C}$ objects.

The minuscule notational difference between (28) and (29$\underline{a}$) makes insufficiently clear the large conceptual divergence intended between these formulas. Although (28) can be subsumed as a special case under (29$\underline{a}$) by allowing $\underline{G}_j$ and $\underline{H}_j$ to include scaling/rescaling functions as well as proper abstractors, the number of component variables in (28)'s micro-output compound $\underset{\wedge}{Y}_j$ is paradigmatically enormous whereas $\underline{G}_j$ paradigmatically comprises just one or a few holistic abstractors that carry configurations of a molar object's module properties at microscopic resolutions revealed e.g. by cellular/molecular biology and atomic physics into the macroscopic features to which our commonsense conceptions of the world are attuned. Similarly, input tuple $\underset{\wedge}{\tilde{Z}}_j$ in (29$\underline{a}$) is paradigmatically a low-dimensional grouping into equivalence classes of what in (28) is a richly multidimensional array $\underset{\wedge}{Z}_j$ of translocated micro-inputs, including micro-indeterminacies that $\underset{\wedge}{\tilde{Z}}_j$ condenses into a small number of molar residuals.

The progression from Def. 2 through equations (27) and (28) to molar behavior (29$\underline{a}$) is logically straightforward. Not so straightforward, however, is the final <u>reduction</u> step this schema makes possible when our original conception of the molar variables at issue does not identify them as composites. To complete reductive explanation of an initially unanalyzed molar regularity

$$\text{In } \Omega, \quad \underline{\check{Y}} = \Psi(\underline{\check{Z}}) ,$$

we must argue first that each object ω in Ω is de re equivalent to a compound object $\underline{o}$ of a kind $\underline{C}$ having a certain structural analysis $\langle \underset{\sim}{P}, \underset{\sim}{L}, \underset{\sim}{R}, \underset{\sim}{k} \rangle$; next, that some law of form (29$\underline{a}$), whose transducer Ψ_j is the same as Ψ, follows from $\langle \underset{\sim}{P}, \underset{\sim}{L}, \underset{\sim}{R}, \underset{\sim}{k} \rangle$ in the manner here described; and finally, that holistically conceived variables $\underset{\wedge}{\check{Y}}$ and $\underset{\wedge}{\check{Z}}$ simply <u>are</u> variables $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{Z}}_j$, respectively. That is, the reduction step contends that for any ω in Ω and $\underline{o}$ in $\underline{C}$ to which ω is de re equivalent, $\ulcorner\underset{\wedge}{\check{Y}};\omega\urcorner$ and $\ulcorner\check{Z};\omega\urcorner$ are really the same molar events as $\ulcorner\tilde{Y}_j;\underline{o}\urcorner$ and $\ulcorner\tilde{Z}_j;\underline{o}\urcorner$, respectively.[19]

Def. 2 and equations (27)-(29a) take no responsibility for this last conclusion. They do, however, show how to derive molar composites of micro-property arrays that may coincide with certain ordinary-thing attributes tightly enough to be plausible candidates for the nature of the latter. Thus, statistical mechanics identifies a molar composite, Mean-kinetic-energy (of a macro-object's particles), whose impressive correlation with phenomenal Temperature across most manipulations of perceived temperature after known factors of perceptual distortion are partialled out urges identifying one with the other. Urges of this sort are seldom resisted in practice; whether they nevertheless ought to be is for philosophers to ponder.

[Macro-dynamics and problems of structural variation.

There is, of course, much more to say about the micro-origins of macro-phenomena than the basic theory of this just sketched. One elaboration, needed for issues of "hierarchical" system organization, is formalizing how analysis of the molar behavior of $\underline{o}$ as a composite of the behavior of its parts at a chosen grain combines with similar reductive analysis for some or all of these parts to derive $\underline{o}$'s behavior from that of its parts' parts, and so on for arbitrarily many iterations of parts/whole analysis. But this is a routine extension of the basic theory that introduces no new conceptual issues beyond the need for hierarchical-organization theory to distinguish clearly between part/whole analytic dependencies and master/slave <u>causal</u> orderings ("command structure") that may or may not be found on a given molarity level.

---

[19] At this point ontological decency demands some concern for the sense in which properties of a tuple can also be viewed as properties of the mereological sum of that tuple's components. The expedient answer is that when ω is de re equivalent to a tuple $\underline{o}$, there must be some parsing function $\mathfrak{z}$ such that $\underline{o} = \mathfrak{z}(\omega)$ while $\underline{o}$'s having a property $\underline{P}$ is equivalent to ω's having not $\underline{P}$ but rather the property signified by predicate '$\mathfrak{z}$(__) has $\underline{P}$'. But I am not confident that this small correction to the equating of $\tilde{Y}_j$ with $\underset{\wedge}{\check{Y}}$ and $\underset{\wedge}{\tilde{Z}}_j$ with $\underset{\wedge}{\check{Z}}$ is all we need here.

A more significant omission from present micro-molar metatheory is system dynamics for compound objects. In principle this too should be routine, insomuch as our previous overview of dynamics (pp. 71-79) allows its system objects to be compounds of any complexity. Nevertheless, problems lurk within the generic model that become obtrusive when we reflect upon the nature of molar abstraction.

Consider basic auto-regressive model (16) rewritten for compound objects of kind C in the notation of (29*a*):

$$\text{In } C,\quad \tilde{Y}f = \Phi(\tilde{Y},\tilde{Z}) \qquad \left(\, \underset{\wedge}{\tilde{Y}} =_{\text{def}} G(\underset{\wedge}{Y}),\quad \underset{\wedge}{\tilde{Z}} =_{\text{def}} H(\underset{\wedge}{Z}) \,\right) \tag{30}$$

wherein each molar variable in $\underset{\wedge}{\tilde{Y}}$ and $\underset{\wedge}{\tilde{Z}}$ is some composite of the micro-variables $\underset{\wedge}{Y}$ and $\underset{\wedge}{Z}$ on whose respective ranges G and H are tuples of abstractors. For this formula to make sense, not only must each object o in C be carried by some excursive displacement f into a unique successor $o' = f(o)$, but also o' must be in the domain of each component of compound variable $\underset{\wedge}{\tilde{Y}}$. Moreover, if the auto-regression is to continue for some subset C' of C, each o' in C' must also be a compound object of kind C. These three conditions are straightforward to realize if each component o_i of each $o = \langle o_1,\ldots,o_n\rangle$ in C has a unique micro-successor $f(o_i)$ such that for any module selector μ on C, $\mu(f(o_1),\ldots,f(o_n))$ is usually of the same causal kind as $\mu(o_1,\ldots,o_n)$. But in fact, almost all commonsense enduring subjects (i.e., macro-things) that can be parsed first of all as a succession of causally linked stages) undergo stage-to-stage alterations in their finer-grained assembly structures, as attested e.g. by the turnover of membership in social communities, the frequent division or sometimes death of individual cells in multicellular organisms, and the incessant interchange of atoms across macro-object boundaries. Real-world applications of model (30) in which $f(o_1,\ldots,o_n) = \langle f(o_1),\ldots,f(o_n)\rangle$ for most $o = \langle o_1,\ldots,o_n\rangle$ in C are simply not to be expected.

However, if the successor $\underline{o}' = \underline{f}(\underline{o})$ of $\underline{C}$-kind object $\underline{o}$ is not also of kind $\underline{C}$, not merely cannot auto-regression (30) be iterated past $\underline{o}'$, but model (30) may well be defeated at the very outset by failure of $\underline{f}(\underline{o})$ even to be in the domain of auto-regressive variable $\underset{\wedge}{\tilde{Y}}$ in the first place. For, each micro-variable composited in $\underset{\wedge}{\tilde{Y}}$ has t-derivational form $\underset{\wedge}{y}_j\mu_k$ for some module-selector μ_k on $\underline{C}$-kind objects. If $\underline{f}(\underline{o})$ lacks some of the assembly features definitive of kind $\underline{C}$, μ_k may not pick out a non-null module of $\underline{f}(\underline{o})$ at all if $\underline{f}(\underline{o})$ contains fewer components than $\underline{o}$, and $\mu_k\underline{f}(\underline{o})$ may not be in the domain of $\underset{\wedge}{y}_j$ even if it exists. For example, suppose that $\langle$John-at-$\underline{t}$, Jane-at-$\underline{t}$, Jimmy-at-$\underline{t}$, Janet-at-$\underline{t}\rangle$ is the cohabiting Smith family--husband, wife, and children--at time $\underline{t}$, but that the successor, Smiths-at-$\underline{t}+\Delta$, of Smiths-at-$\underline{t}$ is only $\langle$Jane-at-$\underline{t}+\Delta$, Jimmy-at-$\underline{t}+\Delta$, Janet-at-$\underline{t}+\Delta\rangle$ because Jane divorced John between $\underline{t}$ and $\underline{t}+\Delta$ and won custody of the children. If module selector μ_4 picks out just the 4th component of tuples for which it is defined, and $\underset{\wedge}{y}$ is any variable defined over person-stages, say IQ, $\underset{\wedge}{y}\mu_4$ has a value for Smiths-at-$\underline{t}$, namely, Janet's IQ at $\underline{t}$, but is undefined for Smiths-at-$\underline{t}+\Delta$ because μ_4(Smiths-at-$\underline{t}+\Delta$) does not have a 4th component even though Janet still exists at $\underline{t}+\Delta$. Alternatively, we could keep Smiths-at-$\underline{t}+\Delta$ a 4-tuple by parsing this as, e.g., $\langle$Jane-at-$\underline{t}+\Delta$, Jimmy-at-$\underline{t}+\Delta$, Janet-above-the-navel-at-$\underline{t}+\Delta$, Janet-below-the-navel-at-$\underline{t}+\Delta\rangle$, but then object μ_4(Smiths-at-$\underline{t}+\Delta$) would not be in the domain of $\underset{\wedge}{y}$ at all much less being of the same sociological kind as μ_4(Smiths-at-$\underline{t}$).

This unstable-structure problem has in principle an abstractly general solution whose formalisms, unhappily, are more complicated than can be made clear in a few words. But since this matter is far too important to ignore, I shall try to sketch its essence with emphasis upon a version that does not explicitly involve dynamics.

The problem of structural variation arises generically for the management of molar regularity as soon as we contemplate how Def. 2 and its abstractive consequence (29$\underline{a}$) might apply to everyday macro-phenomena. To illustrate, suppose that we hope to explain how commonsense generalities of human biology--effects of dietary practices upon health, principles of respiration and blood circulation, etc.--derive from humans' constitutions as cellular organizations. If we start by parsing each human (read human-stage) ω as a tuple o of cells, we find enormous variation in these individuals' cellular numerosities. That is, there is no one integer n such that almost every human is the mereological sum of some n-tuple of cells. Moreover, even were we to select one n and restrict our study just to cellwise-parsed humans $\{o\} = D_n$ that are n-tuples of cells ordered within each o for maximal causal parallel among the n-tuples in D_n (e.g., all bone cells are listed first in each $o \in D_n$, and all cranial cells first among the bone cells, etc.), there will still be so much structural diversity among the μ_k-modules of various o in D_n for any fixed module selector μ_k--e.g., the mean distance or total number of contact points among the cells in $\mu_k(o)$ may differ greatly from one $o \in D_n$ to another--that for few if any causal laws L will $\mu_k(o)$ be in the domain of L for all o in D_n. The point here is simply that even though there may well be a partition $\{D_j\}$ of the class D_* of cellwise-parsed humans within which each particular subclass D_j is a determinately structured compound-object kind whose structural analysis entails a satisfyingly rich array of molar phenomena in D_j (i.e., D_j fixates not only the total numerosity of its members' cells but also their subtype quantities, spacing, connectivities, etc.), the number of humans in any one of these structurally uniform subclasses D_j of D_* is in all likelihood a vanishingly small proportion of D_*. So how might there be molar regularities shared by all or at least a great many of them?

In abstract theory the answer is simple--except that this theory leaves a large gap between the formal solution and its feasibility in practical applications. Suppose that a generic class C_* of structurally diversified tuples does have a partition into species $\{C_j : j \in \mathbf{J}\}$ such that for each j in index set $\mathbf{J}$, compound object-kind C_j has a structural analysis $\langle \mathbf{P}_j, \mathbf{L}_j, \mathbf{R}_j, \mathbf{k}_j \rangle$ under which, for certain input and output tuples $\underset{\wedge}{X}_j$ and $\underset{\wedge}{Y}_j$ of t-derivative micro-variables over domain $\underline{C}_j$, and certain abstractor tuples $\underline{G}_j$ and $\underline{H}_j$ on the ranges of $\underset{\wedge}{Y}_j$ and $\underset{\wedge}{X}_j$ respectively,

$$(31\text{-}\underline{j}:\ \underline{j} \in \underline{J}) \qquad \text{In } \underline{C}_j,\ \ \underline{\tilde{Y}}_j = \Psi_j(\underline{\tilde{X}}_j) \qquad (\ \underset{\wedge}{\tilde{Y}}_j =_{\text{def}} \underline{G}_j(\underset{\wedge}{Y}_j),\ \ \underset{\wedge}{\tilde{X}}_j =_{\text{def}} \underline{H}_j(\underset{\wedge}{X}_j)\ )$$

is a molar law derived from $\langle \underset{\sim}{P}_j, \underset{\sim}{L}_j, \underset{\sim}{R}_j, k_j \rangle$ in the fashion explained earlier for (29$\underline{a}$). But each of these molar laws will prima facie be restricted to just one structural species. For if any component $\underset{\wedge}{\tilde{z}}_j$ of molar variables $[\underset{\wedge}{\tilde{Y}}_j, \underset{\wedge}{\tilde{X}}_j]$ is defined over $\underline{C}_j$ in a way that exploits the assembly structure specific to $\underline{C}_j$, $\underset{\wedge}{\tilde{z}}_j$ is unlikely to include members of a structurally different species $\underline{C}_k$ in its regular domain at all, and may well not have the same causal import for members of $\underline{C}_k$ as it has for $\underline{C}_j$-objects even if it does.

Illustration. Suppose that the assembly structure of $\underline{C}_j$ includes the condition that for each $\underline{o} = \langle \underline{o}_1, \ldots, \underline{o}_{n_j} \rangle$ in $\underline{C}_j$, the parts of $\underline{o}$ that are of a certain numerically scaled micro-type α are exactly $\underline{o}_1, \ldots, \underline{o}_{26}$. And suppose also that for a certain micro-variable $\underset{\wedge}{z}$ whose domain includes all object parts of type α, the first 26 components of t-derivative micro-compound $\underset{\wedge}{Y}_j$ have composition $\underset{\wedge}{y}_{jk} = [\underset{\wedge}{z}\mu_k]$ ($\underline{k} = 1, \ldots, 26$) where $\mu_k(\underline{o})$ picks out just the $\underline{k}$th component $\underline{o}_k$ of $\underline{o} = \langle \underline{o}_1, \ldots, \underline{o}_{n_j} \rangle$. Then if g_{j1} is the function on the range of $\underset{\wedge}{Y}_j$ that abstracts the arithmetic mean of $\underset{\wedge}{Y}_j$'s first 26 components, i.e., if $g_{j1}\underset{\wedge}{Y}_j =_{\text{def}} 26^{-1}(\underset{\wedge}{y}_{j1} + \ldots + \underset{\wedge}{y}_{j26}) + 0\cdot(\underset{\wedge}{y}_{j27} + \ldots + \underset{\wedge}{y}_{jn_j}) = (\sum_{k=1}^{26} \underset{\wedge}{z}\mu_k)/26$, the value of molar variable $\underset{\wedge}{\tilde{y}}_{j1} =_{\text{def}} [g_{j1}\underset{\wedge}{Y}_j]$ for each $\underline{o}$ in $\underline{C}_j$ is the mean value of $\underset{\wedge}{z}$ over all of $\underline{o}$'s type-α parts. But now consider any $\underline{o}' \in \underline{C}_*$ whose specific kind $\underline{C}_k$ ($\neq \underline{C}_j$) does not have the assembly-structural feature just presumed for $\underline{C}_j$. Unless the first 26 components of $\underline{o}'$ are all in the domain of $\underset{\wedge}{z}$, $\underline{o}'$ will not be in the (regular) domain of $\underset{\wedge}{\tilde{y}}_{j1}$ at all. And even if it is, $\underset{\wedge}{\tilde{y}}_{j1}(\underline{o}')$ will still not be the mean of $\underset{\wedge}{z}$ for the type-α parts of $\underline{o}'$ unless $\underline{o}'$ contains exactly 26 of these and they occur first in $\underline{o}'$.

To be sure, we can alternatively define $\underset{\wedge}{\tilde{y}}_{j1}$ in such fashion that its value for any $\underline{o} \in \underline{C}_*$ having at least one type-α part is the mean value of $\underset{\wedge}{z}$ over $\underline{o}$'s type-α parts. (Details do not here matter.) But that still leaves

$\underset{\wedge}{\tilde{y}}_{j1}$ either undefined for $\underline{C}_*$-objects having no type-α parts at all or--if we extend $\underset{\wedge}{\tilde{y}}_{j1}$'s domain to all of $\underline{C}_*$ by fiat--having anomalous values for them.

Even if components of $[\underset{\wedge}{\tilde{Y}}_j, \underset{\wedge}{\tilde{X}}_j]$ are defined in terms of structural properties in $\underline{C}_j$ not shared by objects in $\underline{C}_k$, however, there can still be a component-by-component functional parallel between $[\underset{\wedge}{\tilde{Y}}_j, \underset{\wedge}{\tilde{X}}_j]$ and $[\underset{\wedge}{\tilde{Y}}_k, \underset{\wedge}{\tilde{X}}_k]$ for all $\underline{j}$ and $\underline{k}$ in the index set j for species of genus $\underline{C}_*$. Specifically, suppose for all $\underline{j}, \underline{k}$ in $\underline{J}$ that $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{Y}}_k$ (and similarly for $\underset{\wedge}{\tilde{X}}_j, \underset{\wedge}{\tilde{X}}_k$) not merely have the same range of values and hence the same number $\underline{m}_Y$ of components, but also that for each $\underline{i} = 1, \ldots, \underline{m}_Y$ and each scale value $\underline{\tilde{y}}$ in the common range of $\underset{\wedge}{\tilde{y}}_{ji}$ and $\underset{\wedge}{\tilde{y}}_{ki}$, the property over kind-$\underline{C}_j$ objects represented on the $\underset{\wedge}{\tilde{y}}_{ji}$-scale by $\underline{\tilde{y}}$ is in some yet-to-be-clarified sense functionally equivalent to the property over kind-$\underline{C}_k$ objects represented by $\underline{\tilde{y}}$ on the $\underset{\wedge}{\tilde{y}}_{ki}$-scale.

> <u>Intuitive illustration</u>. Suppose that $\underset{\wedge}{\tilde{y}}_{j1}$ abstracts for each object $\underline{o}$ in $\underline{C}_j$ the mean $\underset{\wedge}{z}$-value of $\underline{o}$'s type-α parts. Then for any other $\underline{k}$ in $\underset{\sim}{j}$, it is intuitive that $\underset{\wedge}{\tilde{y}}_{k1}$ has the same causal import in $\underline{C}_k$ that $\underset{\wedge}{\tilde{y}}_{j1}$ has in $\underline{C}_j$ if $\underset{\wedge}{\tilde{y}}_{k1}$ abstracts the mean $\underset{\wedge}{z}$-value of each $\underline{C}_k$-object's type-α parts, even when $\underline{C}_k$-objects differ greatly from $\underline{C}_j$-objects in how many type-α parts they have. But if $\underset{\wedge}{\tilde{y}}_{k1}$ abstracts the within object variance of $\underset{\wedge}{z}$ over type-α parts, or mean $\underset{\wedge}{z}'$ ($\neq \underset{\wedge}{z}$) over type-$\alpha$ parts, or mean $\underset{\wedge}{z}$ over type-β parts, we would <u>not</u> expect--though we could be proved wrong--that $\underset{\wedge}{\tilde{y}}_{k1}$ plays the same molar role for objects of kind $\underline{C}_k$ that $\underset{\wedge}{\tilde{y}}_{j1}$ does for objects of kind $\underline{C}_j$.

If that is so, we can define molar variables $\underset{\wedge}{\tilde{Y}}_* = [\underset{\wedge}{\tilde{y}}_{*1}, \ldots, \underset{\wedge}{\tilde{y}}_{*m_Y}]$ and $\underset{\wedge}{\tilde{X}}_* = [\underset{\wedge}{\tilde{x}}_{*1}, \ldots, \underset{\wedge}{\tilde{x}}_{*m_X}]$ over the entirety of generic domain $\underline{C}_*$ to be such that for each $\underline{j}$ in $\underset{\sim}{j}$, the restrictions of $\underset{\wedge}{\tilde{Y}}_*$ and $\underset{\wedge}{\tilde{X}}_*$ to $\underline{C}_j$ are $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{X}}_j$, respectively. That is,

$$\text{For each } \underline{o} \text{ in } \underline{C}_*, \text{ if } \underline{o} \in \underline{C}_j \text{ then } [\underset{\wedge}{\tilde{Y}}_*, \underset{\wedge}{\tilde{X}}_*](\underline{o}) =_{\text{def}} [\underset{\wedge}{\tilde{Y}}_j, \underset{\wedge}{\tilde{X}}_j](\underline{o}) \;.$$

Then for each $\underline{j}$ in $\underset{\sim}{j}$, (31-$\underline{j}$) is equivalent to

$$(32\text{-}\underline{j}: \underline{j} \in \underset{\sim}{j}) \qquad \text{In } \underline{C}_j, \quad \underline{\tilde{Y}}_* = \underline{\Psi}_j(\underline{\tilde{X}}_*) \;.$$

Equations (32) still give us a different molar law for each determinately structured $\underline{C}_j$ within $\underline{C}_*$. But these are now just species-differentiated transductions of the same input/output variables; and to unite them under a single transducer for the genus we need only to define (a) a range-$\underset{\sim}{j}$ "structural" variable $\underset{\wedge}{S}$ over $\underline{C}_*$ whose value for any object $\underline{o}$ identifies $\underline{o}$'s structural kind $\underline{C}_j$, and (b) a transducer Ψ_* over the range of $[\underset{\wedge}{\tilde{X}}_*, \underset{\wedge}{S}]$ whose restriction to any particular value $\underline{j}$ of $\underset{\wedge}{S}$ is just Ψ_j. That is,

$$\text{For each } \underline{o} \text{ in } \underline{C}_*, \text{ if } \underline{o} \in \underline{C}_j \text{ then } \underset{\wedge}{S}(\underline{o}) =_{\text{def}} \underline{j} \;,$$

$$\text{For each value } [\underline{\tilde{X}}_*, \underline{j}] \text{ of } [\underset{\wedge}{\tilde{X}}_*, \underset{\wedge}{S}], \quad \Psi_*(\underline{\tilde{X}}, \underline{j}) =_{\text{def}} \Psi_j(\underline{\tilde{X}}_*) \;.$$

(More flexibly, we let each value of $\underset{\wedge}{S}$ demark the disjunction of all $\underline{S}_j$ that are equivalent as input to Ψ_*. And in practice we would seek a multidimensional scaling of $\underset{\wedge}{S}$ that cogently reflects the multiplicity of ways in which assembly structures $\{\underline{S}_j\}$ differ.) Then equations (32) and hence (31) are equivalent to the single molar law

$$\text{(33)} \qquad \text{In } \underline{C}_*, \quad \underline{\tilde{Y}}_* = \Psi_*(\underline{\tilde{X}}_*, \underline{S})$$

whose domain is the entirety of genus $\underline{C}_*$ and which copes with structural variation within $\underline{C}_*$ by treating this as an additional dimension (or dimensions) of input. Each of species-laws (31-$\underline{j}$) is related to (33) as $\underline{L}^{o}$ is to $\underline{L}^{+}$ on p. 82, above; and if we desregard qualms about the ontological status of the variables and transducer in (33), we can claim that (31-$\underline{j}$) derives from and is explained by (33) under the metaprinciple of Strong Domain Constriction.

Elementary but important examples of collating structure-specific molar laws $\{(31\text{-}\underline{j})\}$ into a single generic law (33) are pandemic in statistical sampling theory. When statisticians deduce formulas for the behavior of sample statistics, these are always relative to a fixed sample size $\underline{N}$. (Cf. p. 90ff., above.) And transducers of the molar laws so derived are indeed usually conditional on a particular $\underline{N}$. Yet by matching suitably chosen statistics across sample size

--notably, for any measure $\underset{\wedge}{q}$ over the individual sample members, defining $\underset{\wedge}{\bar{q}}_N$ to be the variable on domain $\underline{D}_N$ of size-$\underline{N}$ samples whose value for each $\underline{s}_N$ in $\underline{D}_N$ is the sum of $\underset{\wedge}{q}$-scores in $\underline{s}_N$ divided by $\underline{N}$, and then collating $\{\underset{\wedge}{\bar{q}}_N\}$ over $\{\underline{D}_N\}$ to yield $\underset{\wedge}{\bar{q}}_*$: Sample-mean-on-$\underset{\wedge}{q}$ without restriction to a particular sample size--we find that these $\underline{N}$-specific statistical laws can usually be subsumed by tolerably well-behaved generic functions in which $\underline{N}$ is a parameter but whose inputs are otherwise sample statistics and population measures prima facie conceived independently of sample size. If you think of $[\underset{\wedge}{\check{Y}}_*, \underset{\wedge}{\check{X}}_*]$ in (33) as an assortment of sample statistics, and structural variable $\underset{\wedge}{S}$ as the dimension of sample size, you will grasp the essence of how SLese overcomes structural diversity, albeit molar systems more integrated than statistical aggregates have vastly more assembly structure to be reckoned with than just numerosity of homogeneous parts.

To extend this construction to the dynamics of macro-systems with unstable assembly structures, we start with $\underline{C}_*$ partitioned among structurally determinate subclasses $\{\underline{C}_j : \underline{j} \in \underset{\sim}{\underline{J}}\}$ as before, but add that for some excursor $\underline{f}$, each $\underline{o}$ in $\underline{C}_*$ has a unique $\underline{f}$-successor that often differs from $\underline{o}$ in assembly structure even though it usually remains in $\underline{C}_*$. We also envision arrays $\{[\underset{\wedge}{Y}_j, \underset{\wedge}{Z}_j] : \underline{j} \in \underset{\sim}{\underline{J}}\}$ of micro-variables respectively defined over these various subclasses $\{\underline{C}_j\}$ of $\underline{C}_*$ in such fashion that whenever a $\underline{C}_*$-object $\underline{o}$ is in $\underline{C}_j$ while its $\underline{f}$-successor is in $\underline{C}_k$ (not necessarily $\underline{k} \neq \underline{j}$), some or all components of compound event $\ulcorner [\underset{\wedge}{Y}_j, \underset{\wedge}{Z}_j]; \underline{o} \urcorner$ are major sources of the various components of compound event $\ulcorner \underset{\wedge}{Y}_k ; \underline{f}(\underline{o}) \urcorner$. Under a sufficiently finer partition $\{\underline{C}_{jk\alpha}\}$ of $\underline{C}_*$ wherein $\underline{C}_{jk\alpha}$ comprises just those

o in C_* for which o is in C_j, its f-successor is in C_k, and $\langle o, f(o) \rangle$ satisfies additional structural conditions α as relevant (e.g., α may impose certain spatiotemporal proximities between components of o and $f(o)$ that are not entailed just by $o \in C_j$ and $f(o) \in C_k$), we should have that for each $C_{jk\alpha}$ in this partition there is some compound micro-law

$$\text{In } C_{jk\alpha}, \quad Y_k f \;=\; \Phi_{jk\alpha}(Y_j, Z_j, E_{jk\alpha})$$

wherein $\underset{\wedge}{E}_{jk\alpha}$ is a tuple of residual inputs in principle specific to the assembly structure of $C_{jk\alpha}$. From there, judicious applications of molar abstractors to these micro-laws, followed by the same formal tricks that synthesize (33) out of (31), gives us a molar auto-regression

$$\text{(34)} \qquad \text{In } C_*, \quad \tilde{Y}_* f \;=\; \Psi_*(\tilde{Y}_*, \tilde{Z}_*, \tilde{E}_*, S_*)$$

wherein $\underset{\wedge}{S}_*$ is a structural variable whose value for any o in C_* identifies the particular $C_{jk\alpha}$ to which o belongs (or better, an equivalence class thereof). Moreover, while $\underset{\wedge}{S}_*(o)$ codes the assembly structure not merely of o but of o's developmental extension $\langle o, f(o) \rangle$, if structural change throughout the development of enduring C_*-things occurs lawfully, $\ulcorner \underset{\wedge}{S}_*; o \urcorner$ is determined in C_* by the assembly structure $\underset{\wedge}{S}(o)$ just of o together with other conditions in or near o prior to the coming about of $f(o)$'s structuring. That is, variable $\underset{\wedge}{S}_*$ should be replaceable in (34) by a function just of $\underset{\wedge}{S}$ and other factors that, unlike $\underset{\wedge}{S}_*$, do not translocate the structure of o's f-successor into a pretend property of o.

I fear that this sketch of molar dynamics is far too compressed to be very perspicuous, but its technicalities do not really matter here. The important point is that causal recursions among molar properties of coarsely classified macro-objects appear feasible so long as (*a*) we can establish abstract trans-structural equivalences under which certain local variables $\{\underset{\wedge}{x}_j\}$ respectively defined over disjoint assembly structures $\{C_j\}$ become usefully collated as domain restrictions of a global variable $\underset{\wedge}{x}$ over the union of $\{\underline{C}_j\}$, and (*b*) we can treat features of local assembly structures as causal properties whose groupings into variables act like additional input/output dimensions in molar causal laws. Accordingly, it is noteworthy that commonsense conceptions of macro-properties are indeed often fraught with concern for their bearers' micro-structures, either positively by requiring particular assembly features for their correct ascription as in (*b*), or negatively by combining properties that are constituted in part by a determinate assembly structure into structure-liberated disjunctions as in (*a*). Conceptions of the sizes, shapes, and spatial locations of molar objects nicely illustrate the first case; conceptions of statistical properties that cut across sample size--i.e., our being able to think about sample means, variances, correlations, etc. without explicitly conditionalizing these on a particular *N*--are paradigm examples of the second. Even so, these practices raise deeply challenging questions for at least the philosophy of science and perhaps for its applied methodology as well.

Regarding (*b*), insomuch as the structural relations among input/output loci in micro-causal laws appear to play a very different role in the ontology of causal determination than do those micro-properties of the input loci that make the output event come out one way rather than another, molar predicates whose significance is appreciably but not entirely structural may greatly complicate if not confute efforts to develop theories of molar causality that are formally isomorphic to models that seem appropriate for the bottom level (if there be such) of micro-causality. Similarly it is unclear whether we can always tell coherent causal stories in which variables whose values are features of assembly structure (e.g. size and shape) are treated as though they are causal inputs/outputs. But an articulate theory of molar explanatory ordering is needed to give these apprehensions purchase. For now, the salient admonition is that when we seek to analyze what it is for unparsed molar object ω to have some holistically conceived property P, we must expect that α's P-hood may be as much a matter of α's having a certain organization of parts as it is of these parts having one array of local properties rather than another.

As for (*a*), the very poverty of preconditions under which equations (32) follow from equations (31) signals that molar models (33) and (34) must have practicality complications not yet acknowledged. For although I have tried to make collation of $\{\underset{\wedge}{\tilde{Y}}_j\}$ into $\underset{\wedge}{\tilde{Y}}_*$ (and similarly $\{\underset{\wedge}{\tilde{X}}_j\}$ into $\underset{\wedge}{\tilde{X}}_*$) intuitively palatable by declaring that for each scale-value $\underline{\tilde{Y}}$ common to local variables $\underset{\wedge}{\tilde{Y}}_j$ and $\underset{\wedge}{\tilde{Y}}_k$, *having-value-*$\underline{\tilde{Y}}$*-of-*$\underset{\wedge}{\tilde{Y}}_j$ is to be in some sense the same property-tuple over $\underline{C}_j$-kind objects that *having-value-*$\underline{\tilde{Y}}$*-of-*$\underset{\wedge}{\tilde{Y}}_k$ is over objects of kind $\underline{C}_k$, this heurism plays no role in construction of (32/33) from (31). Rather, the construction's formal nature is simply the following: Let $\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m$ be any variables whose respective domains $\underline{D}_1,\ldots,\underline{D}_m$ are disjoint and whose respective ranges have all been put into one-one correspondence by some index set $\underset{\sim}{q}$, i.e., for each $\underline{i} = 1,\ldots,\underline{m}$, $\{\underline{x}_{iq}: \underline{q} \in \underset{\sim}{q}\}$ is a listing of the alternative values possible on variable $\underset{\wedge}{x}_i$. Then we can fuse $\underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m$ into a single range-$\underset{\sim}{q}$ variable $\underset{\wedge}{x}_*$ whose domain $\underline{D}_*$ is the union of $\underline{D}_1,\ldots,\underline{D}_m$ simply by stipulating that for each value $\underline{q}$ of $\underset{\wedge}{x}_*$, *having-value-*q*-of-*$\underset{\wedge}{x}_* =_{def}$ *having-value-*$\underline{x}_{1q}$*-of-*$\underset{\wedge}{x}_1$*-or-having-value* $\underline{x}_{2q}$*-of-*$\underset{\wedge}{x}_2$*-or-*...*-or-having-value-*$\underline{x}_{mq}$-of-$\underset{\wedge}{x}_m$. This disjunctive collation works for *any* choice of domain-disjoint local variables with isomorphic ranges under *any* index coordinations thereof, and so can be expected to yield a scientifically *useful* variable $\underset{\wedge}{x}_*$ only under special constraints to which the abstract collation is blind. Clearly, the practical effectiveness of molar models (33) and (34) cannot be indifferent to which local abstractions we choose to equate over different assembly structures. But *how* these choices matter, and what rational should guide them, still want ajudication.

In abstract generality, the collation constraint needed for (33/34) to be practical seems straightforward enough: *The transducer* Ψ_* *under which the* *collated input dimensions* $\underset{\wedge}{\tilde{X}}_*$ *of conjectured generic law* (33) *determine its collated output* $\underset{\wedge}{\tilde{Y}}_*$ *must be inductively accessible*. (Similarly for (34).) That is, by observing scores on enough components of $[\underset{\wedge}{\tilde{Y}}_*, \underset{\wedge}{\tilde{X}}_*, \underset{\wedge}{S}]$ (or on sufficiently good estimators of them) in humanly finite samples of objects drawn from generic domain $\underline{C}_*$, we must be able to estimate Ψ_* with enough accuracy that Est(Ψ_*) applied to our knowledge or decent estimate of $[\underset{\wedge}{\tilde{X}}_*, \underset{\wedge}{S}](\underline{o})$ for additional objects $\underline{o}$ in $\underline{C}_*$ will yield a useful estimate of $\underset{\wedge}{\tilde{Y}}_*(\underline{o})$ even when $\underline{o}$'s values on $\underset{\wedge}{\tilde{X}}_*$ and $\underset{\wedge}{S}$ are at some remove from any input combinations that occur in the samples from which we have inferred Est(Ψ_*). Such inductive accessibility is above all the operational essence of a genuine functional law (see p. 31f., above). It seems evident that were the micro-variables $[\underset{\wedge}{Y}_j, \underset{\wedge}{X}_j]$ and abstractors $\langle \underline{G}_j, \underline{H}_j \rangle$ generating molar regularity (31-$\underline{j}$) for $\underline{C}_j$-structured objects to be chosen independently for each $\underline{j} \in$ _, only by an extraordinary fluke could array $\{(31\text{-}\underline{j}) : \underline{j} \in \underset{\sim}{j}\}$ of local phenomena have any trans-species patterning under which our identifying the local transducers $\{\Psi_j : \underline{j} \in \underset{\sim}{j}'\}$ in some subset $\underset{\sim}{j}'$ of $\underset{\sim}{j}$ tells us anything about Ψ_k for any $\underline{k}$ in the remainder of $\underset{\sim}{j}$. But conversely, regardless of how our conception of collated variables $[\underset{\wedge}{\tilde{Y}}_*, \underset{\wedge}{\tilde{X}}_*]$ may have arisen, if we do seem able to learn the entirety of Ψ_* by induction from finite sample data (nevermind how we judge our prospects at this), then we can have no better reason for taking the compound molar property *having-value*-$\langle \underline{\tilde{Y}}, \underline{\tilde{X}} \rangle$-*of*-$[\underset{\wedge}{\tilde{Y}}_j, \underset{\wedge}{\tilde{X}}_j]$ scaled in one $\underline{C}_*$-species $\underline{C}_j$ by $[\underset{\wedge}{\tilde{Y}}_*, \underset{\wedge}{\tilde{X}}_*]$-value $\langle \underline{\tilde{Y}}, \underline{\tilde{X}} \rangle$ to be component-by-component causally equivalent to the compound property *having-value*-$\langle \underline{\tilde{Y}}, \underline{\tilde{X}} \rangle$-*of*-$[\underset{\wedge}{\tilde{Y}}_k, \underset{\wedge}{\tilde{X}}_k]$ scaled by $\langle \underline{\tilde{Y}}, \underline{\tilde{X}} \rangle$ in any other $\underline{C}_*$-species $\underline{C}_k$.

Of course, this epistemic induction criterion for trans-structural equivalence of molar properties offers little operational advice for initiating such collations. But free choices in this respect seldom arise in practice. In those special cases where we begin with known or conjectured micro-laws and seek to infer molar behavior from their compoundings, notably, in data-analytic applications of sampling theory, it is usually evident what must be put with what in order to write conceptually well-behaved generic functions in which assembly structure sets parameters. (E.g., no one at all familiar with elementary statistics would dream of collating the mean of variable $\underset{\wedge}{x}$ in odd-sized samples with $\underset{\wedge}{x}$'s variance in even-sized samples.) And in those far-more-prevalent research situations where molar events unparsed for micro-composition are what we most directly perceive or infer from our instrument readings, causally equivalent properties are already collated for us across structural kinds by interface mechanisms through whose mediation these act as input to our cognitions. Thus in the vastly oversimplified but philosophically standard example of inference to dispositions, when we identify ρ-ability as whatever property enables certain objects to react $\underline{R}$-wise when tested on input $\underline{T}$, only at more advanced stages of inquiry into $\underline{T}/\underline{R}$ couplings are we sometimes able to conclude that the structural composition of ρ-ability in objects of one kind differs from its composition in certain other objects that are also ρ-able.

This doing-what-comes-naturally selection of micro-variables $\{\underset{\wedge}{Y}_j, \underset{\wedge}{X}_j\}$ and abstractors $\{\underline{G}_j, \underline{H}_j\}$ to constitute molar variables $[\underset{\wedge}{\tilde{Y}}_*, \underset{\wedge}{\tilde{X}}_*]$ in laws that generalize over structural kinds largely obviates our need to craft an applied methodology for this. It does not, however, insure that these natural selections work well. Even if molar variables $\underset{\wedge}{\tilde{Y}}_*$ and $\underset{\wedge}{\tilde{X}}_*$ cannot be improved upon in the epistemic cogency of their micro-constitutions and trans-structural collations in generic domain $\underline{C}_*$, this may still not give us inductive access to generic transducer Ψ_* in (33) over the entirety of $\underline{C}_*$. Thus for a particular $\underline{C}_*$-species $\underline{C}_j$, some component $\underset{\wedge}{\tilde{x}}_j$ of $\underset{\wedge}{\tilde{X}}_j$ may encorporate assembly features of $\underline{C}_j$ that have no significant counterpart in the assembly structures of other $\underline{C}_*$ -species $\underline{C}_k$, thereby thwarting collation of $\underset{\wedge}{\tilde{x}}_j$ with any molar dimension $\underset{\wedge}{\tilde{x}}_k$ over $\underline{C}_k$ whose sample-estimated behavior in $\underline{C}_k$ is inductively diagnostic of $\underset{\wedge}{\tilde{x}}_j$'s behavior in $\underline{C}_j$. (E.g., if $\underset{\wedge}{\tilde{x}}_j$ is the mean of type-α components within each $\underline{o}$ in $\underline{C}_j$, there may well be no useful counterpart of $\underset{\wedge}{\tilde{x}}_j$ over objects having no type-α components.) We can always extend any such $\underset{\wedge}{\tilde{x}}_j$ into an $\underset{\wedge}{\tilde{x}}_*$ whose domain includes all of $\underline{C}_*$ simply by creating an anomolous $\underset{\wedge}{\tilde{x}}_*$-value that holds for any object lacking assembly features required for inclusion in $\underset{\wedge}{\tilde{x}}_j$'s regular domain. But then what we learn about generic transducer Ψ_* in (33) from $\underline{C}_*$-samples that include only regular values on $\underset{\wedge}{\tilde{x}}_*$ will inductively imply little or nothing about what $\Psi_*(\underline{\tilde{X}}_*,\underline{S})$ may be for inputs in which $\underline{\tilde{X}}_*$ contains the anomolous $\underset{\wedge}{\tilde{x}}_*$-value. Thus, while construction of (33) from (31) is always formally possible, (33)'s practicalities are best captured by its fractionation into

$$(35\text{-}\underline{k}:\ \underline{k}\in\underset{\sim}{k})\qquad \text{In } \underline{C}^*_k,\quad \underline{\tilde{Y}}^*_k = \Psi^*_k(\underline{\tilde{X}}^*_k,\underline{S})\ ,$$

wherein $\{\underline{C}^*_k:\ \underline{k}\in\underset{\sim}{k}\}$ is a partition of $\underline{C}_*$ into structural regions, coarser than $\{\underline{C}_j:\ \underline{j}\in\underset{\sim}{j}\}$ (i.e. each $\underline{C}^*_k$ generally includes many different specific structural kinds $\underline{C}_j$), under which each $[\underset{\wedge}{\tilde{Y}}^*_k,\underset{\wedge}{\tilde{X}}^*_k]$ contains only components of $[\underset{\wedge}{\tilde{Y}}_*,\underset{\wedge}{\tilde{X}}_*]$ whose values in $\underline{C}^*_k$ are all regular and, for as many $\underline{k}\in\underset{\sim}{k}$ as possible, regional transducer Ψ^*_k is inductively accessible from feasible sampling of $\underline{C}^*_k$.

Were partition $\{\underline{C}^*_k\}$ of $\underline{C}_*$ to comprise only a small number of regions whose molar regularities can be discerned separately with equal ease, the difference between array (35) and its formal unification in (33) would be only a minor technicality. But in practice, regional laws $\{(35\text{-}\underline{k})\}$ may vary enormously in the success with which we can learn them: For some $\underline{k}$ in $\underset{\sim}{k}$, Ψ^*_k may be a mathematically docile function connecting molar variables $[\underset{\wedge}{\tilde{Y}}^*_k,\underset{\wedge}{\tilde{X}}^*_k]$ that comprise a small number of components readily identifiable (excepting residuals) by trained perception or standard instrument interpretations. Whereas for other $\underline{k}$ in $\underset{\sim}{k}$ (in all likelihood the far more prevalent ones), Ψ^*_k may be so disorderly that we could scarcely make sense of it even were the components of $[\underset{\wedge}{\tilde{Y}}^*_k,\underset{\wedge}{\tilde{X}}^*_k]$ not so esoteric in their abstractive derivations and collations that we can scarcely conceive them in the first place. This point has already been well-illustrated by our

Shadows example (p. 45ff.). For enclosure-stages $\{v\text{-at-}t\}$ whose assembly structures insure that descriptors such as 'the point-like major light source in ___', 'the shadow of the protrusive object contained in ___', and '___'s opaque bounding planes' have well-defined referents when applied to v-at-t, it is relatively straightforward to work out laws relating dimensions of shadow size to light-source positioning and other commonsensically ascertainable features of enclosure v at time t. But let falter the structural preconditions for these descriptors to succeed, and not only do our molar shadow laws fail, we are left with little notion of what variables might effectly describe and explain the patterning of light and darkness within enclosures of uncontrolled contents.

A somber practical corollary here is that when we seek to work out molar regularities within and between developmental stages of enduring macro-things with evolving assembly structures, we are fortunate indeed if we can find much <u>stability</u> (see p. 48f.) therein. For in (35)'s counterpart for molar dynamics, even if the f-successors of most C_*-objects are also in C_*, we may still have that when o is in one structural region C^*_k, $f(o)$ often shifts to another within which molar regularity is most naturally characterized by a rather different selection of abstracta. As a result, whereas disclosure of system dynamics is a standard and feasible aim of advanced research in the more molecular sciences, the prospect of success at this becomes increasingly precarious as we ascend into higher levels of molar abstraction unless--as is true of computers and many other engineered systems but <u>not</u> of most natural ones--our enduring subjects have essentially invariant assembly structures.]

<u>The future of reduction-SLese.</u>

The theory of micro-molar explanation sketched above is just that--a <u>theory</u>. More precisely, Def. 2 and equations (27)-(35) exhibit a framework for reductive explanation of macro-phenomena that derives with deductive immediacy from modern science's standard formalisms for expressing lawfulness. But formalisms need to be fleshed out with some determinate semantic content before they actually say anything, and the history of ideas is littered with conceptual schemata that have gone nowhere. How often, or in what ways, can we expect this framework--call it "reduction-SLese"--to have useful payoff in real applications? I suggest that its current prospects are at least sixfold:

First of all, as already emphasized, modern sampling theory and multivariate data analysis are built upon instantiations of reduction-SLese that for better or worse have become pandemic in the behavioral and biological sciences. Admittedly, these methodologies have evolved in part under their own inner urges, with resulting relevance for empirical research often more tenuous than is entirely defensible. But that is a powerful practical reason for seeking to tease out, with perspicuity and depth, just what views of mathematics and nature shape current data-analysis methodology, how well these are aligned with prevailing goals of substantive research in the sciences they serve, and whether consanguinity within the methodology/substance marriage cannot be enhanced by a clearer delineation of each partner's compatible but imperfectly coincident positions in a common space of SLese concerns. For example, data of the sort treated by Multidimensional Scaling research generally have a formal organization that cannot happily be subsumed under the classical Variables-by-Subjects or Variables-by-Subjects-by-Occasions design formats to which almost all the multivariate-methodology literature has remained restricted. Multidimensional scaling models do, however, fit nicely into SLese's more general micro-molar framework in a fashion that makes evident their essential continuity with older orthodoxies.

Secondly, applied physics, chemistry, and engineering have demonstrated remarkable success in deriving the macro-properties of well-structured compound objects from those of their parts. Most artifacts of modern living--automobiles, radios, calculators, etc., not to mention numerous crafted materials of which most of us are scarcely aware--work in the consumer-relevant ways they do mainly because someone knowledgeable of certain micro-causal regularities designed a felicitous assembly of parts locally governed by those laws. (Not all the relevant micro-laws, or their macro-consequences under a particular assembly design, are known in advance, however--which is why applied engineering includes a large dollup of cut-and-try.) I am insufficiently conversant with engineering technologies to claim with confidence that all or even many of these practices can be effectively verbalized in SLese; but if not, then surely it will be instructive to identify respects in which they find reduction-SLese wanting.

Thirdly, the working contents of sciences such as sociology, economics, and macro-biology, whose primary objects of concern are evidently local groups of some sort, abound in large fragments of reduction-SLese. For, most of these disciplines' basic variables are explicitly conceived as composites of micro-properties distributed within each macro-object (local group) in the variable's domain. So any lawful relation proposed to govern these variables realizes the molar side (i.e. (29a), (33), or (34)) of SLese's micro-molar model. Usually absent from these accounts, however, is the articulate micro-story envisioned by Def. 2. Sometimes this may be because the micro-laws at issue seem too commonsensical to need detailing. Thus in studies of predator/prey population dynamics, there is prima facie little mystery in how, for any living mammal s at time t, s's wellbeing at time $t+\Delta$ is determined inter alia by s's species, ambient temperature, quantities of nearby water and vegetation, and encounters between t and $t+\Delta$ with other animals of various species. (Even then, commonsensical as this may seem, actually writing out micro-laws of local survival is a task of great challenge.) But for whatever reason, the micro-events from which group phenomena abstract are seldom conceived in terms precise enough for their lawfulness to be characterized in SLese. For example, in sociological study of how Homocide Rate is affected by Handgun Incidence in a macro-domain of urban communities, it is easy to operationalize indices of Homocide Rate in terms of documents available in police files; but what are the micro-variables over what micro-domains whose values are supposedly diagnosed by these records? Presumably, certain human interactions count as "homocides." But is Homocide then a binary variable over the domain of all person-stage pairs such that for each <person-s_h-at-time-t_i, person-s_j-at-time-t_k>, s_h-at-t_i either kills or does not kill s_j-at-t_k, while Handgun Incidence abstracts from the Smallarmedness variable that sees each s-at-t as either possessing or not possessing a handgun? Or are there better ways to define the relevant micro-variables? The point here is not that working out a micro-model of homocide should be especially difficult, but simply that not until we *have* worked this out can we embed the molar dependency of Homocide Rate on Handgun Incidence in a full-blooded SLese reduction that explains the molar patterning.

There is, of course, no obligation for local-group studies to accompany their molar findings with explicit micro-reductions for these. We know in general from sampling theory how molar compositing manages to abstract order from the seeming chaos of micro-events; and if we are content with regularities already identified in congenial macro-terms, wallowing in their underlying micro-complexities may seem pointlessly masochistic. Even so, there can be little doubt that the group properties studied by sociology, economics, etc. do indeed analytically abstract in reduction-SLese fashion from properties and relations among individual group members. Explicating these particular reductions should be relatively straightforward; and learning how to cope with the unanticipated difficulties that will arise if we actually attempt to carry some of these through should greatly enhance our capability for reductive analysis of molar phenomena whose variables' a-derivational character is far more obscure.

Fourthly, there are areas of scientific inquiry wherein, for a given class C_* of macro-objects, much has been learned or conjectured about the local behavior of C_*-object modules at a fine grain of parts-parsing while a considerable holistic lore of $\underline{C}_*$-things also exists, yet significant linkage between these two outlooks on $\underline{C}_*$ has yet to be achieved. In such cases, the reduction-SLese framework provides guidance for exploring alternatives in which aspects of the one approach may or may not reduce to the other. Psychology's paramount instance of such double vision is human neurophysiology vs. classical and neo-classical (cognitive-science) accounts of mentation; and in Part II we shall examine the SLese prospects of bringing these together.

Fifthly, recent philosophy of mind has given increasing prominence to the thesis of *psychological functionalism*. Under the slogan that mental attributes are "functional states," this contends roughly that mental predicates signify roles in higher organisms' inner causal processes at the abstraction level of "machine-table states" in computational systems. The literature on this theme has advanced to a level of considerable technicality (see Block, 1980) --which is as it should be, except that its sophistication is misshapenly unbalanced. Its philosophical esoterica are legitimate enough, but have been grounded upon conceptions of causal mechanism so primitive that these premises/arguments/conclusions have dubious relevance to more realistic models of how machines and organisms work. Although there is much to commend in the functionalist outlook, none of its proposals can be taken seriously until its explication of "functional state" recognizes the causal/compositional complexity of macrosystems in at least the detail here formalized by Def. 2 and equations (29*a*)-(35).

Functionalism's need is not merely to appreciate the SLese-articulated structure of natural systems; even more crucial is sensitivity to the ontological puzzles that protrude from this at almost every turn. Especially germane are issues of molar causality, notably, whether it is possible to tell coherent causal stories in which abstracta and their bases figure jointly or, if not, how systems of causal laws can be stratified into molarity levels within each of which the anti-symmetry of explanation is preserved. Similarly salient are the differences, or lack thereof, ($\underline{a}$) between an event $\ulcorner \underset{\wedge}{x};\underline{o}'\urcorner$ and its translocation $\ulcorner \underset{\wedge}{x}\underline{f};\underline{o}\urcorner$ when $\underline{o}' = \underline{f}(\underline{o})$ (see Note 3 p. 84a), ($\underline{b}$) between the properties of an unparsed molar object ω and of a tuple $\underline{o}$ de re equivalent to ω (see fns. 15 & 19), and ($\underline{c}$) between a molar variable $\underset{\wedge}{\tilde{y}}_j$ over some structurally determinate domain $\underline{C}_j$ and the restriction to $\underline{C}_j$ of the generic molar variable $\underset{\wedge}{\tilde{y}}_*$ defined by collation of $\underset{\wedge}{\tilde{y}}_j$ with more-or-less-comparable local variables $\{\underset{\wedge}{\tilde{y}}_{j'}\}$ over assembly structures $\{\underline{C}_{j'}\}$ alternative to $\underline{C}_j$ (see p. 119f.). These and related questions, which all evoke the generic problem of what supervenient causal roles we can usefully conceive as played by what constructed entities, are central to psychological functionalism; and although reduction-SLese does not reveal their answers, it does make them accessible in the context of hard science doing real work.

Finally, reduction-SLese establishes an advanced base camp for philosophic exploration of reduction/emergence. It is a matter of deductive logic that if $\underline{o} = \langle \underline{o}_1,\ldots,\underline{o}_n \rangle$ is any tuple of an object ω's parts, and $\underline{o}$ is of compound-object kind $\underline{C}$, then <u>some</u> properties and behaviors of ω are supervenient upon $\underline{o}$'s microstructure in the fashion formalized by equations (27)-(29). But whether everything worth saying about ω can be so analyzed is another question altogether. It has often been argued that wholes are more than mere aggregates of their parts; and it is certainly true that listing the elementwise properties of ω's proper parts does not suffice to entail the entirety of ω's holistic character. But compound object $\underline{o}$ is not just an aggregate; it also has an assembly structure in which its part/part relations are just as essential to $\underline{o}$'s causal kind as are its parts' nonrelational properties. For example, if $\underline{o} = \langle \underline{o}_1,\ldots,\underline{o}_n \rangle$ is a disassembled furniture kit at time $\underline{t}$ whose successor $\underline{f}(\underline{o}) = \langle \underline{f}(\underline{o}_1),\ldots,\underline{f}(\underline{o}_n) \rangle$ at time $\underline{t}+\Delta$ has been reorganized into a chair, the kit-at-$\underline{t}$ and chair-at-$\underline{t}+\Delta$ are nothing more, as objects, than $\underline{n}$-tuples $\underline{o}$ and $\underline{f}(\underline{o})$, respectively; and moreover each part $\underline{f}(\underline{o}_i)$ of chair $\underline{f}(\underline{o})$ has essentially the same nonrelational properties as its predecessor o_i in the kit. But $f(o)$, unlike o, is an actualized *chair*--not through some emergentist miracle but simply because the spatial relations of $f(o)$'s parts give it a stable assembly structure with useful molar properties that the kit lacks.

I am myself still far from convinced that all properties of macro-objects, in particular all that are in some sense "structured," can be equated with abstracta from structural analyses of tuples to which the object is de re equivalent. But provisionally, lacking any clear evidence to the contrary, that's the way to bet.