William F. Lawvere
*State University of New York, Buffalo*

## EXPLICIT FOUNDATIONAL CONCEPTS IN THE TEACHING OF MATHEMATICS

*[As the organizers requested, this article has a non-technical component. I share the motivation of that request, since it reflects a desire to convey to the general reader a sense of the importance of foundational questions. On the other hand, my space is limited, so I have mainly sketched expositions which teachers with some background will be able to develop. I suggest that readers of this book should become acquainted with the rudiments of category theory, and I hope that the glimpses I am able to give here will provide further motivation for such a study.]*

Mathematics education is closely linked with the foundation of mathematics; only the teaching and learning process can transmit, to those who will apply it, the foundation that has been concentrated in any given epoch. The design of foundation is thus strongly influenced by the needs of education. An outstanding example is Richard Dedekind's struggle 150 years ago to find the clearest way to explain differential calculus to his students at the Zurich Politechnicum, a struggle which resulted in his famous analysis of the continuum, now an icon of foundations. Some more recent developments that are claimed to have foundational importance, such as elementary topos theory, were also strongly influenced in their origin by the need to explain calculus. Teaching can incite individual scientists to discover – and make explicit – concepts that before had been implicit in collective thinking, and in turn, the explicit concepts accelerate learning, as well as research.

The principle "avoid concepts" is proposed as pedagogical doctrine in the US and in other countries, justified by the pragmatic claims that concepts do not lead immediately to the needed applications or that concepts are too difficult to teach. In this note I want to oppose these mystifications.

Explicit concepts, definite enough to be the basis of reasoning, are indispensable to applications and to education. Though it is clear that the concepts that are concentrated from applications may sometimes need testing to evaluate their consistency, and it is also clear that much work is involved in presenting concepts in such a way that they can be understood by all the students who consciously participate in learning, such presentations need not be burdened with the historical constructions that have been used to establish consistency. For example, a perfectly rigorous treatment of infinite-dimensional differential geometry (that is, of the mathematics underlying the everyday physics of continuous bodies and waves), does not have to be preceded by a long detour into topological vector spaces or countable additivity, nor does it require an elaborate arbitrary build-up of atlases of charts, based in turn on rituals devised by Bolzano and Cauchy, based in turn on a cumulative hierarchy of sets. All those steps might be useful in developing some particular aspect of the subject, but they should not be a barrier that prevents people from learning to work rigorously.

I want to discuss four examples of concepts that can be made to seem "too advanced to teach" by a culture that exploits the unstated assumption that the above pyramidal sort of foundation is the only rigorous one. Even those teachers who may not agree with the thesis of this article concerning the origin of foundations, may find that the clear introduction of these concepts permits a freedom and a confidence for further advance that is not possible when the concepts are left unmentioned.

(1) A concept that can usefully be made explicit early in school is that of homomorphism. There are basic examples, such as angle measurement and Napier's map from multiplication to addition. The concept of homomorphism is helpful in understanding why the importance of Napier's achievement remains undiminished despite the disappearance of the slide rule. More generally, there is a very useful principle that a transformation of systems of intensive quantities corresponds (contra-variantly) to a smooth transformation of their domain spaces if and only if the transformation satisfies a few algebraic tautologies. This principle is a theorem [9] in so many varied categories that in applications it can be taken as axiomatic in its own right; but it is also intuitive: for example, if one comes upon a probability distribution for which every random variable has standard deviation zero, there should be a point at which the whole distribution is concentrated.

Precise reasoning about the world – and also intuitive reasoning – involves frequent replacement of intensively variable quantities by constants. That replacement is effected by distributions, or extensive quantities. Inaccuracy in such reasoning can result from failing to recognize these concepts and thus failing to sufficiently investigate the variety of distributions. This variety can be limited by the further axiom that defines the notion of averaging process: if the variable quantity happens to be constant, then its expectation is the same constant; but there is typically still an infinity of such averaging processes. Below we will return to a frequent paradigm that generates concrete notions of distribution, but first I want to emphasize the simplicity of generality.

(2) A concept which needs to be made explicit for everyone is the concept of functional. The claim that it is "too advanced" may find favor with those who have endured pyramidally-founded courses; in some context all the material in those courses is required, but almost none of it is necessary for the basic concept of functional. That concept is just this: in the dependence of the one quantity z on another variable quantity f, a whole panorama of information within f may be required to determine one value of z (for example, the behavior z at this moment of a sword – or a horse shoe – depends on the whole history f of its tempering by a craftsman). The concept of functional can become the basis of rigorous reasoning if we isolate the notion of a map with some specified target Z,

$$Y^X \dashrightarrow Z$$

but whose source $Y^X$ is a function space (or perhaps a subspace of a function space.) The meaning of function space is uniquely determined by the tautology that maps

$$A \longrightarrow Y^X$$

from any space A are in natural bijection with maps

$$X \times A \longrightarrow Y$$

a bijection that has been fundamental since the earliest days of the calculus of variations. This precise meaning of functional and this uniqueness of the structure of function space holds within any given category in which the function spaces exist. There are many such categories and greatly simplified methods for constructing them have been developed in recent decades. Much basic reasoning is independent of the category one might choose and indeed, as I showed in my thesis [5], the assumption of the function space functor can be used as a direct axiom on a category, with powerful consequences.

Around the time when the concept of functional was coming to be recognized by Ascoli, Volterra [8], and others, and after over 150 years of investigation – in the calculus of variations – of functionals of real- and vector-valued functions, the functionals of truth-valued functions were isolated and incorporated into the logical calculus under the name of existential and universal quantifiers.

(3) A concept that should be explicit after 300 years of differential calculus is the concept of tangent, which expresses in a precise way the principle that, in the infinitely small, curves become straight. To every smooth space X is associated the smooth space $X^T$ of all tangents to X, and this relation is functorial in the sense that for every smooth map $X \xrightarrow{f} Y$, there is an induced smooth map

$$X^T \xrightarrow{f^T} Y^T$$

and if $Y \xrightarrow{g} Z$ is another smooth map, then

$$(gf)^T = g^T f^T$$

for the composite maps. That functoriality is the chain rule of differential calculus. There is a natural map $X^T \longrightarrow X$ indicating for each tangent vector its base point of origin and hence, for each point x, there is the subspace $T_x$ of $X^T$ consisting of all tangents based at x. For any "curved" map f, if fx = y, then the restricted map

$$T_x \xrightarrow{f_x} T_y$$

is "straight" in that it satisfies the homomorphic property

$$f_x(\lambda v) = \lambda f_x(v)$$

for all natural stretchings $\gamma$ of the tangent concept. (In a smooth category, this one homomorphic property implies all the other properties commonly associated with straightness, at least for a dense subcategory of tangent spaces.)

There are many functors, but of course the basis of differential calculus is the special nature of the tangent functor in particular: if judiciously chosen, some of its special properties serve well as axioms for direct application. Yoneda and Grothendieck [3] emphasized forty years ago an important method for analyzing particular functors, exploiting the representability of some functors. Representability means that there is a single space T which concentrates the essence of the whole functor; more precisely, the functor is (up to isomorphism) just the process $(\ )^T$ of forming function spaces with arbitrary target spaces, but fixed source T. A lemma of Yoneda implies that if a functor is representable, then the representing space is unique, so that the study of properties of that space is an important support for the study of the functor; for example, the natural operations $\lambda$ on the functor can all be represented by endomorphisms of the representing space. It is fortunately relatively easy to construct categories that contain ordinary space-time, etc., but where the tangent bundle functor becomes representable; in other words, the tangents themselves are just infinitesimal paths. (This justifies my choice of notation $(\ )^{T}$ for the tangent bundle functor). Simultaneously, one obtains a universal formula

$$(Y^X)^T \cong (Y^T)^X$$

for the tangent space of any infinite-dimensional function space. That formula is rather complicated to prove if the space T is not mentioned. It is fortunate that no nonstandard maps between (or points of) the ordinary spaces are introduced by this construction, and also fortunate that T itself is quite simple: it has just one point and its function algebra is only two-dimensional.

The recognition of the representability of the tangent functor had been advocated in the past, with varying degrees of explicitness, by Leibniz, Study, and Kähler, and is now common in the theory of algebraic groups [1], though not yet in the calculus of variations and in functional analysis. The resulting role of nilpotent quantities is quite compatible with standard practice in physics and engineering, but in the past, that observation had remained underdeveloped, partly because keeping track of the various quantities can appear complex without the recognition of the principle that every map has a definite target. That principle was forced on category theory by algebraic topology, where the behavior of the Poincaré and Hurewicz functors would be incomprehensible without it. (There is a frequent necessity to change the source and target, but this is effected by composing with other maps, yielding new maps with possibly different properties.)

(4) I wish to emphasize the simplicity and fundamental nature of the concept of extensive quantity. Already Grassmann [2] had lamented the tendency to recognize only intensive quantities, thus denying extensive quantities their just status. Subjective idealist philosophers might claim that extensive quantity does not exist, while objective idealist philosophers (like Hegel) might claim that there is no difference. The simple fact is that quantities such as mass, charge, volume, energy, entropy, and probability are fundamentally extensive rather than intensive, and the contrast is still used by teachers of thermodynamics. Foundations of mathematics should lend support to such aspects of teaching.

Ubiquitous in mathematics is a pair of modes of variation of quantity (which of course get combined in complex situations), namely covariant and contravariant dependence on the domain space. On the one hand, contravariant dependence implies such features of intensive quantity as

(a) the inclusion of constant quantities into each algebra of variable quantities;

(b) the possibility of evaluating any variable quantity at any point of its domain of variation; and

(c) the typical problem of extending a variable quantity to a larger domain.

On the other hand, none of those features apply in general to extensive quantities. The extensive mode of variation has features shared by any covariant functor E on a category of spaces. (For nontriviality, assume that $E(1) \neq 0$ where 1 is the one-point space): The covariant functoriality itself means that for each smooth map X ⟶

Y there is an induced "push-forward" map $EX \longrightarrow EY$, and that composite maps induce composites; hence there is the support problem of determining whether a given extensive quantity in E(Y) is supported on $X \longrightarrow Y$ in the sense that it comes from some extensive quantity in E(X). For example, the distribution of whales on the earth Y comes from the distribution of whales in the ocean X, hence the former is "supported" on the inclusion of X into Y.

There are many problems in combinatorial probability that gifted students solve "intuitively"; but a useful gift to all students would be the recognition that the main step in the solution is often the analysis of a distribution of interest as the image along an appropriate map of a distribution on some other space X, where this antecedent distribution is uniform, i.e. invariant under all automorphisms of X. Try to calculate the distribution of results of rolling a pair of dice, where the crucial ingredient is a summation map to the 11-point space Y from a 36-point space X.

Any covariant functor E will have special relation to the one-point space 1. That space is characterized by the property that any space X has a unique map $X \longrightarrow 1$ to it. The functoriality thus induces

$$E(X) \xrightarrow{\int_X} E(1)$$

which can be so denoted because it assigns to each distribution in X its total. The water in the clouds above a city has a complicated distribution but, in particular, has a total. Experts will recognize that the usual distributions of compact support constitute a good example of such a functor E, but so do measures, in a different category. On the other hand, the maps $1 \longrightarrow X$ are the points x of X and applying the functoriality of E, we obtain for each point x a map

$$E(1) \xrightarrow{\delta_x} E(X)$$

that assigns to each possible total a distribution on X that has that total, but which is supported on x. Thus the dreaded Dirac delta distributions are actually a tautology of great generality when their extensive nature is recognized; if one tries instead to present them as generalized intensive quantities, the resulting mysteries and paradoxes can block the students' progress.

As hinted above, one can go quite far with model-building on the basis of simple direct axioms about types E of extensive variation of quantity, but one does need explicit examples. A general kind of example that might be called the Riesz paradigm is available in any category having the exponential (= function space) construction, namely, given two spaces R and V and a system $\Lambda$ that operates on both, define

$$E(X) = \mathrm{Hom}_\Lambda(R^X, V)$$

as the subspace of $V^{(R^X)}$ consisting of all those $\varphi$ (ways of reducing R-valued intensively-variable quantities f to V-valued constants) that satisfy the condition

$$\varphi[\lambda f] = \lambda\varphi[f]$$

for all $\lambda$ in $\Lambda$. Often one chooses R and $\Lambda$ so that $E(1) = V$, in which case the Dirac delta is concretely realized as evaluation and provides an inclusion $X \longrightarrow EX$.

On the other hand, if $V = R$ and if $\Lambda$ is chosen quite large, for example $\Lambda = R^R$, then the intuition of variables with zero variance strongly suggests that $X \longrightarrow EX$ should be an isomorphism for suitable V, i.e. that the only V-valued extensive quantities that are so highly homogeneous, are just the point-evaluations. The only spaces X for which such a principle is not true have so far come from the world of measurable cardinals, but that world is demonstrably quite remote from the world of ordinary pure and applied mathematics. Our foundation should take notice of this fact.

There are of course other concepts (decisive abstract general relations) that compassionate teachers can present to students in a way that helps to illuminate their path. A small system of concepts – consistently chosen – serves directly the everyday learning, development and use of mathematics; moreover, the principles that characterize these concepts can in fact be shown to logically imply all of known mathematics.

*REFERENCES*

[1] Demazure M., Gabriel P., *Groupes algébriques*, North Holland, 1970.

[2] Grassmann H., *Ausdehnungslehre*, 1844. Reprinted Chelsea, N.Y. 1969.

[3] Grothendieck A., *Famille d'espaces complexes et fondaments de la géométrie analytique*, Séminaire Henri Cartan 13e année, 1960/61, exposé 11. Secrétariat Mathématique, Paris 1962.

[4] Lawvere F.W., Schanuel S.H., *Conceptual Mathematics, a first introduction to categories*. Cambridge University Press, 1997

[5] Lawvere F.W., *The category of categories as a foundation for mathematics, La Jolla conference on categorical algebra*, Springer-Verlag, 1966, 1-20.

[6] Toposes of Laws of Motion, 1997, video transcript http://www.acsu/buffalo.edu~wlawvere/list.html

[7] *Categorie e Spazio, un profilo* Lettera Matematica Pristem 31, Springer Milano, 1999, 35-50.

[8] Volterra's functionals and the covariant cohesion of space, in R. Betti (editore), *Categorical Studies in Italy*, Supplemento ai Rendiconti del Circolo Matematico di Palermo, Serie II, 64, 2000, 201-214.

[9] Stone M.H., *The generalized Weierstrass approximation theorem*, Math. Mag. 21, 1948, 167-184.