30 June 1972
Vol. 176, No. 4042

LETTERS
An Emperor's Work: T. Komai; Space Shuttle Costs: M. Bader; Animal Experiments: W. U. Gardner

EDITORIAL
Are the Data Worth Owning?: J. R. Macdonald

ARTICLES
Plate Tectonics and Australasian Paleobiogeography: P. H. Raven and D. I. Axelrod
The Tornado Threat: Coping Styles of the North and South: J. H. Sims and D. D. Baumann
Information Technology: Its Social Potential: E. B. Parker and D. A. Dunn

NEWS AND COMMENT
Anabolic Steroids: Doctors Denounce Them, but Athletes Aren't Listening
NCI Announces Plans for Fort Detrick
Dual Publication: "Ingelfinger Rule" Debated by Scientists and Press
Rainmaking: Stockholm Stand Watered Down for Military
Nobelists: Piccioni Lawsuit Raises Questions About the 1959 Prize

RESEARCH NEWS
Chalones: Chemical Regulation of Cell Division

BOOK REVIEWS
The Biology of Lampreys, reviewed by J. W. Atz; The Biology of the Cell Cycle, J. E. Sisken; Thermodynamic Theory of Structure, Stability and Fluctuations, R. Zwanzig; Volcanoes, R. E. Stoiber

REPORTS
Early Thai Bronze: Analysis and New Dates: D. T. Bayard
Potential Energy Surface Including Electron Correlation for F + H₂ → FH + H: Refined Linear Surface: C. F. Bender et al.
Retrograde Axonal Transport in the Central Nervous System: J. H. LaVail and M. M. LaVail
Viral RNA Subunits in Cells Transformed by RNA Tumor Viruses:  
N. Tsuchida, M. S. Robin, M. Green ......................... 1419

Spontaneous Transformation of Human Brain Cells Grown in vitro and  
Description of Associated Virus Particles: J. Hooks et al. .............. 1420

Salicylate: A Structure-Activity Study of its Effects on Membrane  
Permeability: H. Levian and J. L. Barker .......................... 1423

Endosperm Protein Synthesis in Maize Mutants with Increased Lysine Content:  
P. S. Misra et al................................................... 1425

Biosynthesis of Hemoglobin Ann Arbor: Evidence for Catabolic and Feedback  
Regulation: J. G. Adams III et al. .................................. 1427

Somatic Cell Genetic Assignment of Peptidase C and the Rh Linkage Group to  
Chromosome A-1 in Man: F. Ruddle et al. .............................. 1429

Cytochalasin B: Inhibition of d-2-Deoxyglucose Transport into Leukocytes and  
Fibroblasts: S. H. Zigmond and J. G. Hirsch .......................... 1432

Histological Changes in Lobsters (Homarus americanus) Exposed to Yellow  
Phosphorus: D. E. Aiken and E. H. Byard ............................ 1434

Mixed Function Oxidase and Ethanol Metabolism in Perfused Rat Liver:  
I. E. Hassinen and R. H. Ylikahri .................................. 1435

Wiswesser Line Notation: Simplified Techniques for Converting Chemical  
Structures to WLN: D. A. Koniver, W. J. Wiswesser, E. Usdin  .......... 1437

Pollinators in High-Elevation Ecosystems: Relative Effectiveness of  
Birds and Bees: R. W. Cruden ...................................... 1439

Calcium and Sodium Contributions to Regenerative Responses in the Embryonic  
Excitable Cell Membrane: S. Miyazaki, K. Takahashi, K. Tsuda ............. 1441

Behavioral Thermoregulation by Fishes: A New Experimental Approach:  

Rearing Complexity Affects Branching of Dendrites in the Visual Cortex of  
the Rat: F. R. Volkmar and W. T. Greenough .......................... 1445

Target-Set and Response-Set Interaction: Implications for Models of Human  
Information Processing: H. Egeth, N. Marcus, W. Bevan ................ 1447

Technical Comments: Theta Rhythm and Memory: W. R. Klemm; P. W. Landfield,  
J. L. McGaugh, R. J. Tusa ....................................... 1449
Are the Data Worth Owning?

Amitai Etzioni has recently (14 April 1972, page 121) raised the question of who should ultimately own the data. He points out that, since data is (sic) often lost or becomes inaccessible, agencies that finance data collection or preparation should require that it be made available to others by deposition in a data bank or library.

A more fundamental question than who should own the data is, Are the data even worth owning? Unfortunately, the answer is usually an embarrassing and costly “No” across the entire spectrum of research. The problem usually lies in lack of knowledge about the trustworthiness of the data. Measures of uncertainty are usually not given at all; even when they are, they are themselves generally untrustworthy. Lancelot Hobgen has stated that “less than one percent of research workers clearly apprehend the rationale of statistical techniques they commonly invoke.”

Further, David Lide, head of the National Standard Reference Data System of the National Bureau of Standards, estimates that from 50 to over 90 percent of the published raw data available for producing trustworthy, evaluated results for the physical properties of scientific materials cannot, in fact, be used for this important purpose. A good illustration of the basic difficulty has been given by the late W. J. Youden of NBS. He states that, of 15 observations of the mean distance to the sun published from 1895 to 1961, each worker’s estimated value is outside the uncertainty limits set by his immediate predecessor.

Both systematic and random errors occur in all experimental situations. They should be estimated, discussed, and cited separately, as Churchill Eisenhart has pointed out. Ideally, systematic errors should be estimated by independently measuring the quantity in question with a different apparatus, preferably one that operates on a different principle from that of the original apparatus. One should strive to make the estimated maximum systematic error comparable to or smaller than the estimated root-mean-square random error of the experiment. When it is impractical to obtain independent estimates of the systematic error, a good rule of thumb is to multiply one’s best estimate of it by a factor of 3.

Even when estimates of individual errors (deviations) are calculated, it is unusual for the experimenter to check the deviations for stochastic independence and to state the result of such a check. Rarely indeed does one find mention of the statistical distribution that the deviations appear to follow. Without such knowledge, however, one cannot assess the meaningfulness of such important derived quantities as ordinary confidence limits and standard deviations. When individual measurements are to be analyzed by such techniques as least squares, one seldom finds experimenters replicating the individual measurements closely enough to obtain trustworthy estimates of uncertainties for use in weighted least squares. Yet only can one verify the assumptions implicit in even unweighted least squares. Nonlinear least squares analysis is becoming much more common these days, but one never finds it shown that the bias in parameter estimates introduced by this estimation technique is safely smaller than the sampling error. Finally, there are almost always random errors present in the values of all variables measured (except in whole number cases), not just in the “dependent” variable, as is usually assumed in ordinary least squares analysis. Although a generalized least squares technique is necessary and available to handle such situations, it is hardly ever used—nor is the need for it usually recognized.

Clearly, much further education in data analysis, presentation of results, and the need to call in a statistician is necessary before a high proportion of published data can be properly used for more than qualitative purposes.—J. Ross MacDonald, Chairman, Numerical Data Advisory Board, National Research Council.