# Approaching Error in Scientific Knowledge and Science Education

PAUL ZACHOS[1], ROBERT PRUZEK[2], THOMAS HICK[3]

*[1]ACASE, 110 Spring St., Saratoga Springs, NY 12866 paz@acase.org, [2]State University of New York at Albany, Education 236, 1400 Washington Ave., Albany, NY 12222, rmpruzek@yahoo.com; [3]retired, New York State Department of Education, 223 Seward Pl., Schenectady, NY 12305, thick@nycap.rr.com*

**ABSTRACT**. The search for dependable knowledge is a basic human concern. An inescapable aspect of this search is the risk of error in making judgments and claims. All aspects of scientific investigation from initial conceptualization to final presentation of findings are fraught with possible error. Ideally, scientific methodology may be seen as a formalization of techniques that have evolved historically to provide scientific investigators with protection against various sources of error through control and counteraction. Explicit, constructive, and even creative approaches to dealing with error often constitute hallmarks of successful scientific research. Still, the role of error in the conduct of scientific research, and in the accumulation of scientific understanding is usually not part of science education. How should the concern for the quality and dependability of scientific knowledge make its way into science education curriculum? At what points in curriculum should the topics of error, and methods for dealing with error, be introduced? Findings from developmental psychology, cognitive science, and science education research can provide suggestions for the design of curricula that will help science students develop sensitivity, understanding, and competence in dealing constructively with the role of error in the development of dependable knowledge about the natural world.

## 1. Introduction

In describing the critical moment when Johannes Kepler dismissed his hypothesis of circular planetary orbits around an off center sun, Richard Feynman (1965) illustrates a fundamental feature of scientific decision making. Kepler's decision followed his discovery that the position of the planet Mars predicted by his model differed by 8 minutes of arc from the observations of his predecessor Tycho Brahe. His confidence in the accuracy of Tycho's observations was so great that he concluded that this difference was too great to be ignored. Because of the level of precision of his own predictions and his confidence in the known reliability of Tycho's work he felt confident in rejecting his previously held theory, and moved to other ones, leading finally to the theory of elliptical planetary orbits. Kepler had faith in the exacting and precise work of his fellow scientist to the degree that he was ready to reject his own cherished model as being 'in error' and enter again into the struggle of building a model for the orbit of Mars.

Our own interest in the subject of error began with a study of the effects of competence in specified scientific inquiry capabilities on success in discovering concepts in experiential settings. In one task, secondary school students were given sets of cubes of varying sizes and densities, and three beakers of liquids also of varying densities, such that a wide range of possibilities for floating and sinking were available. The students were also provided with the laboratory equipment needed to build and test hypotheses about why things float or sink. We found consistent correlations in these experiments between the ability to distinguish potential sources of error, and success in building and testing models of the phenomenon under consideration (Zachos et al. 2000).

This paper is a preliminary foray into a vast field. It is directed to identifying a place for the issue of scientific error in the secondary school science curriculum. To do so, we will specify and define several features of error to be addressed, indicate why they are central to science education, and then, for each, provide an example drawn on the one hand from the history and philosophy of science, and on the other from cognitive and developmental research in the area of science education.

## 2. The Importance of Error and the Role of Error in Science

The subsuming thesis in Karl Popper's discourses on the building of scientific knowledge in *Conjectures and Refutations* (1963) is that we can use error to learn. The role of error in the building of scientific knowledge could not be more fundamental than it is for Popper: 'There is no criterion of truth at our disposal, and this supports pessimism. But we do possess criteria which, if we are lucky, may allow us to recognize error and falsity' (p. 37).

Attempts to take a comprehensive look at sources of error which affect knowledge of the natural world go back as early as the work of Francis Bacon. In *Novum Organum*, Bacon (1620) identifies 'idols' or properties of human perception, personality, and enculturation which constitute biases that prevent us from achieving a 'true interpretation of nature'. He even defines a category of idols whose origin lies in the ways that we conduct science.

In *Error and the Growth of Experimental Knowledge*, Deborah Mayo (1996) surveys how experimental scientists go about learning from mistakes. She also posits a central role for error in the building of scientific knowledge. She calls for greater attention to the characteristics and effects of error:

> The view that we learn form error, while commonplace has been little explored in philosophy of science. When philosophers of science do speak of learning from error—most notably in the work of Popper—they generally mean simply that when a hypothesis is put to the test of experiment and fails, we reject it and attempt to replace it with another. Little is said about what the different types of errors are, what specifically is learned when an error is recognized, how we locate precisely what is at fault, how our ability to detect and correct errors grows, and how this growth is related to the growth of scientific knowledge. (p. xvii)

## 3. The Scope of This Paper

In this paper, our attention is focused on the role that error plays when human intention is directed to the acquisition of systematic knowledge of the world. More specifically, our discussion aims at the scientific enterprise and to the corresponding educational enterprise, in so far as the latter is directed to the acquisition of scientific knowledge and skill via the relationship between teacher and student. Moreover, we will only look at two aspects of error in science, those having to do with indeterminacy and bias. We therefore exclude broad notions of error such as the one proposed by Ernst Mach (1905) who construes error in terms of behaviors that disadvantage a species within the grand scheme of adaptation and evolution. Also, we do not emphasize errors that take the form of misconceptions, or misinterpretations of observations or of phenomena, despite their centrality to a larger foray into scientific errors. Our attention is mostly limited to 'claims' of fact in the context of conveying to students a sound understanding of scientific principles and knowledge, and threats to those claims. The term 'claim' is used to refer to any statement that asserts scientific backing for its veracity. This includes facts, principles, rules, laws, and interpretations of these. Errors are common to all aspects of scientific investigation – design, conceptualization, observation, experimentation, data collection, analysis, and presentation. And it is essential, we believe, that students of science understand the role of error in the discovery of scientific knowledge, as well as the need for healthy skepticism, about all scientific claims. We consider it important that students understand the roles of error in the conduct of science, and how naïve acceptance of so-called scientific 'facts' can undermine a healthy appreciation for the accomplishments of scientists.

## 4. Two Uses of the Term Error in Science

There are two senses of the word 'error', each representing threats to the validity of scientific claims, that will be considered here:

- Error associated with 'indeterminacy'

- Error associated with 'bias'

### 4.1. ERROR ASSOCIATED WITH INDETERMINACY

There are two related, but distinctive kinds of error associated with indeterminacy. The first, concerns 'measurement error', and the second, 'sampling error'. Measurements of the same object or phenomenon often vary between definable limits rather than being unique and absolute. As a consequence, a single definitive value is often difficult to obtain when trying to characterize an object of phenomenon. Cases in point are atomic weights of elements, and measuring the speed of light, which was first done effectively by Michelson in 1922, but for which many refinements followed in later in the 20th century. Variations can be influenced by the precision of the measurement instrument, properties of the object, control of the process of observation, and characteristics of the persons or methods making the measurements (*e.g.* fatigue); indeed, 'measurement error' is usually considered to have 'chance-like' or 'unsystematic' properties. One of the simplest and most effective ways to deal with this type of variation is to take repeated measures of the same object (e.g. location of a celestial object) and to calculate the average of the obtained values. 'The point of averaging over many repetitions is to reduce the sampling error in the average value to a negligible amount' (Mosteller 1968). This chance-like variation is usually referred to as 'measurement error' and it has been the subject of much study, especially in sciences where measurement problems are often associated with scientific study of phenomena. Quantification of the amount or degree of measurement error is usually characterized as the analysis of reliability (cf. Taylor 1982).

'Sampling error' is of a different kind. This is the error that occurs when one uses an incomplete 'sample' of observations to make statements about a larger domain, usually called a population. The most common form of this kind occurs when sample statistics, like a 'sample mean', are used to make 'inferences' about a population parameter, a 'population mean'. Inferences of this kind generally entail 'sampling error', since samples are by definition subsets of populations, and therefore, contain insufficient information to make wholly accurate or sound statements about population characteristics. Failure to attend to 'sampling error' can lead to misinterpretations of data. Analysis of sampling variation, its sources, and quantification, are part of the field of statistical inference.

Through the exercise of experimental controls and data analysis one can attempt to isolate and quantify the degree to which variation has chance-like properties (Cochran 1968), including both 'measurement and sampling error'. The discipline of inferential statistics takes a step beyond simply modeling 'sampling error', and provides means for characterizing the degree of indeterminacy and uncertainty in one's observations and claims. Feynman holds that it is just in this ability to characterize uncertainty in one's claims that one is being scientific as opposed to the practice of speaking in unconditional terms about what is true and false, possible and impossible (Feynman 1965).

## 4.2. ERROR ASSOCIATED WITH BIAS

Unlike 'measurement error', errors associated with 'bias' are systematic, not random. 'Bias' refers to a distortion of results when a set of values for some object or thing being measured are influenced by some other variable or factor not used in the calculation of scores for the variable itself. For example, if girls tend to systematically interpret questions on a test differently from boys (or any group, in relation to another) then scores derived from that test may tend to be biased in a fashion that corresponds to the ways that student subgroups corresponded to the questions. 'Bias' can be described as the systematic, non-random, effect of unintended and often unanticipated factors distinct from the one(s) under direct observation. The effect of 'bias' is systematic rather than chance-like because it sends values derived from investigations off in a particular direction either in support or opposition to the effect of interest, but in either case suggesting the presence or absence of an effect in a misleading way. Being able to identify sources of 'bias' for a particular field of inquiry may require years of immersion in that discipline (Mosteller 1968); however, there are some sources of 'bias' that are common across all fields.

A source of bias that has been documented among scientists as well as the general public is called 'confirmation bias'. This is the tendency to be predisposed to search for evidence that will support one's own theory. Cognitive research shows that the tendency to favor information that supports one's current hypothesis, theory, or belief (i.e., one's scientific claim), and to disfavor information that challenges these is widespread and persistent (Kuhn et al. 1988). This form of 'bias' is associated with the tendency to be constrained in one's observations by the framework of one's theory (Martin 1970). Darwin was cautious to protect himself from 'bias' by maintaining a separate notebook in which he recorded observations that were incongruent with his theory, thereby helping him to avoid overlooking or underestimating them (Kaplan 1964).

'Bias' can also be associated with convenience or self interest which can be no less threatening to the scientific investigations and presentations:

> We may consciously or unconsciously judge the pointer of an instrument to be on a line on the scale because it makes it easier to take a reading or confirms a previous reading or makes the graph more obviously a straight line. (Fairbrother et al. 1997)

The modern discipline of experimental design pioneered by R.A. Fisher (1935) provided systematic procedures for neutralizing the effects of selection bias, primarily through 'randomization'. Experimental design since his time has relied on the notion that random assignment of experimental subjects to treatments can control for the effects of bias as well as to provide a basis for distinguishing the relative contributions of experimental factors and sampling error. More recently, statistical techniques have been developed to control for selection bias in situations where true experiments (i.e. those that use random assignments) are not possible. The latter kind of study is properly called an observational study. Methods such as Propensity Score Analysis (Rosenbaum 1995) have been developed to reduce the effects of selection biases in observational studies, so as to get closer to the situation that obtains when randomization is used to form treatment groups.

The scientific enterprise generally operates under the assumption that error will surface, and in the course of time will make itself evident. It is assumed that if a claim is inconsistent with nature, that this inconsistency will manifest as a discrepancy between observation and claims, or as an internal logical inconsistency within or between claims. The assumption is not one of blind faith, however. Scientists set up procedures that facilitate the discovery of error and help deal with it when this happens. Criteria and procedures established historically by scientists can be seen in the light of the possibility that they will help ensure opportunities to discover errors. For example, requirements for inter-subjective agreement, the double-blind experiment, or the confirmation of observations by independent observers increases the likelihood of discovering an error of observation or conceptualization (Kaplan 1964). This can also be seen as the reason behind the demand for extreme explicitness in scientific investigations—'In Science … it is useful to publish every bit of empirical evidence, even every conjecture; indeed, no scientific edifice should be built until the plan and materials of its structure have been widely known, judged, and sifted' (Goethe 1792).

How the confrontation with error leads to progress in scientific methods can be characterized by the case of the development of the method of least squares by Andre Legendre. Legendre was called in to deal with questions of irregularities in both the phenomenon and observations associated with the work of the astronomers Pierre-Francoise-André Méchain and Jean-Baptiste-Joseph Delambre. Méchain and Delambre had attempted to define a standard unit of length (i.e. the meter) as one ten-millionth of the distance from the North Pole to the equator by extrapolation from the Paris meridian distance running from Dunkirk to Barcelona (Alder 2002). Legendre devised the method of fitting a curve by minimizing the deviations of squared values from the curve in order to deal with these problems. The technique, for which Karl Fredrick Gauss is also often given historical credit, has since become a foundation for many inferential statistical methods, and underlies many of the methods used both in the experimental design tradition of Ronald Fisher and the observational study tradition characterized by Rosenbaum (1995).

## 5. Error and Science Education

Based on their experiences teaching error analysis in the physical sciences, Séré, Journeaux, and Larcher (1993) conducted a study of first-year Paris University students' conceptions of measurement as they carried out their laboratory work in optics and electricity. The optics tasks involved two methods for measuring the focal length of converging lenses; the electricity task was one of assessing resistance by measuring length, diameter, and resistance of metal wires. Students were asked to obtain and evaluate measures, and to construct confidence intervals for their measurements. Séré and her colleagues found that in spite of instruction in measurement techniques and error analysis, students failed to distinguish random and systematic errors, and that few had critical insight into the notion of a confidence interval. While students maintained that 'the more measurements the better', they did not understand why this is the case. Lubben and Millar (1996), concentrating on students' understanding of reliability of measured data, found that while the use of statistical techniques increased with age in a large sample of students aged 11, 14, and 16, that there was little improvement in understanding what makes one measure better than another. In our own work, both in and out of school settings, we found that most secondary school students who conducted inquiries to isolate the causal factor underlying the period of the pendulum did not take into account sources of systematic or sampling error in interpreting the values they derived from measurements, and failed to spontaneously make use of averaging in spite of years of familiarity with the procedure through math courses (Zachos et al. 2000; Zachos, 2003).

What are we to infer from this picture? Is conceptualizing and dealing with error, uncertainty, and probability too challenging a subject matter for secondary school students? For entry-level university students? Are the theoretical underpinnings of error, uncertainty, and probability too demanding, or does our science education fail to provide the experiences and direction needed when measurement, observation, and empirical inference are taught in secondary schools?

Piaget & Inhelder's (1975) research into children's conceptions of chance suggests that a developmental factor may be at play in the development of competence in dealing with uncertainty. They suggest that understanding and competent action related to application of methods to deal with indeterminacy are dependent upon a number of factors. These include a 'search for order and its causes' and the attainment of the notion of 'reversibility' of logical operations as part of logical-mathematical representation of the world. It is in the nature of conventional logical-mathematical operations to be reversible through the properties of inverses, reciprocals, and negation. For example, the measurement of the distance from A to B logically corresponds to the measurement of the distance from B to A. However, the presence of chance-like perturbations in measurement on the one hand, or bias on the other can lead to irreversibility. It is the very nature of explicit, logical, and precise claims that they create conditions of irreversibility. Thus, Piaget and Inhelder hold that the confrontation with indeterminacy will follow from having formed and attempted to apply logical-mathematical models of the world. Polanyi (1958) suggests that the adoption and increasing refinement of explicit logical-mathematical models of phenomena may itself be closely intertwined with the confrontation with error. He proposes that moving to a state of the 'articulateness' (such as might be associated with logical-mathematical representations of phenomena) initially puts us in a state of fumbling and uneasiness regarding our relationship to the world, in that it does not adequately represent our 'tacit' grasp of the world and makes the enterprise increasingly subject to error:

> Although the gains made by casting our thoughts into articulate terms eventually outweigh by far these initial disadvantages, there will always remain certain chances of error—and even of grave error—which arise from our very adoption of an articulate interpretative framework. This risk is therefore inherent in the exercise of all higher forms of human reason…When superstition is superseded by philosophy and theology, or by mathematics and natural science, we become involved once more in new systems of fallacies from which our practice of mathematics, science, philosophy, or theology can never be strictly free. The mind which entrusts itself to the operation of symbols acquires an intellectual tool of boundless power, but its use makes the mind liable to perils the range of which seems also unlimited. The gap between the tacit and the articulate tends

> to produce everywhere a cleavage between sound common sense and dubious sophistication…(pp. 93-94)

We have seen how success in confronting the problem of indeterminacy is construed by Feynman as a criterion for scientific competence. We have seen that techniques originally developed to deal with error evolve into methods that have become some of the basic workhorses of scientific investigation. Polanyi's analysis suggests how the scientific enterprise, to the degree to which it is characterized by explicit logical-mathematical modeling of the natural world, invites the phenomenon of error as a worthy opponent and takes advantage of its corrosive properties as a tool for moving to increasingly powerful conceptualizations of phenomena. One can see that error is natural and can be useful and informative. In fact, modern curricula reflect an interest in incorporating error and bias into science curricula, as exemplified at: http://www.edu.gov.on.ca/eng/document/curricul/secondary/science/scieful.html (cf. Ministry of Education & Training, Ontario, Canada: 1999). This highlights the need on the part of student and teachers to attend to the phenomenon of error.

## 6. Intended Learning Outcomes – Envisioning a Curriculum Related to Error in Science

We have attempted to make the case that concern for the effects of error and the intentional confrontation with error are central features of the scientific enterprise, and that recognizing and dealing with the phenomenon of error is a worthy goal for science education. Error presents a threat to knowledge. For this very reason, cognizance of error and competence in dealing with it constitutes a criterion for scientific literacy. The question of when and how to develop capabilities in this domain deserves research, and we propose that the research take into consideration the potential value of the following learning objectives as part of a science curriculum that deals with error.

SENSITIVITY AND UNDERSTANDING

- Ways to Conceptualize Uncertainty (e.g. variation in measurement)
- Chance-like vs. Systematic Variation
- Distributions as Summaries of Values
- Distributions as Representations of Phenomena
- Precision vs. Accuracy
- Bias
- Intersubjectivity

SKILLS

- Summarizing Measures (e.g. central tendency and spread, stem and leaf, graphing of data)
- Characterizing Degrees of Certainty and Uncertainty of Values (e.g. the language of probability)
- Sampling
- Randomization
- Generating and Articulating Causal Inferences

- Analyzing Causal Inferences to Test Their Validity (e.g. with venn diagrams, symbolic logic, truth tables)
- Understanding or coordinating the level of measurement precision necessary to display the hypothesized phenomenon.

In exploring and validating such a curriculum, we propose that instructional programs be considered that involve building and testing logical or mathematical models in the context of direct investigations into natural or manufactured phenomena. Keeping things as simple as possible, while remaining close to 'accepted facts' about nature or phenomena, seems especially important. We believe that the teacher's knowledge of the history of science, including known mistakes, errors, and misadventures, can be especially helpful in the context of helping students to develop sound concepts about phenomena; moreover, students will profit if their instruction includes information about the role of errors, uncertainties, chance, and indeterminacy. We advocate that instruction not be confined to presentation of just 'final' or 'accepted truths' about phenomena. In such experiential contexts the 'search for order and its causes' can be elicited and assessed. When effectively structured, direct inquiries into natural phenomena can take advantage of the fact that nature responds to human action, giving direct, and often immediate, feedback as to the efficacy of concepts and methods. In working directly with natural phenomena, uncertainty will always be present and opportunities for introduction of the concept of bias in observation and analysis can be provided 'in situ'.

## 7. The Dark Side of Error

We have presented a picture that suggests a positive role for error and the benefits to be associated with a study of error; but there are problematic considerations as well. For many persons it can be disconcerting to confront the idea that errors are not only possible, but in some domains of science, commonplace. It is common for human beings to seek the security of scientific information, facts, or judgments from experts. The notion that the world about us is a reality of which our thinking may sometimes provide only feeble representations can be perplexing and distressing to some. Moreover, scientists, and those who rely on science for their authority, may not like to admit the incompleteness of or limitations to their knowledge. Indeed, there may be a social stigma associated with error that is threatening to scientist, teacher, and student alike. Science classrooms do not generally function in the light of the positive value of error. For the most part, the concepts presented in textbooks or identified in state curricula are treated in classrooms as 'the truth', and the arbiter of truth becomes the classroom test or high stakes examination. It is our hope that attention to the positive side of error can begin to exert a counterbalance to this force.

## 8. Next Steps

This initial essay is intended to serve as a prelude to more extensive examinations of the role of error in science and science education. There are additional aspects of error in science that have direct implications for curriculum and instruction. We have not been able to give attention to the extensive research literature related to what are called 'misconceptions' in science and mathematics (Novak 1987; Confrey 1990; Roschelle 1998) and to the studies of errors in inference and logical fallacies (Bady 1979; Park et al. 2001). The conceptual change literature, in particular, points to many positive roles of 'misconceptions'. Student errors can serve as anchor points for motivating students and targeting instruction (Fisher & Lipson 1986). Fisher and Lipson's recommendations indicate the value of having 'a science of student error analysis'. In turn, this could have a salutary effect on the common responses to 'failure' by the educational community, and lead to more positive and constructive approaches to student error in the educational mainstream that could be helpful to students.

We intend to pursue further the contribution and implications of the discipline that historically has most explicitly and intentionally confronted the issue of error—the disciplines of statistics and experimental design. In this field, the role of error becomes formal and intentional, even strategic. Error is conceptualized and modeled both mathematically and graphically. Clear distinctions are made between different sources of error, for example, errors that may be reduced through actions of the experimenter by contrast with those that are out of the experimenter's range of control and various sources of bias in selection. Specific techniques have been developed to address these different sources of error, such as the use of chance-like distributions in modeling and hypothesis testing. The experimental/statistical discipline provides principles and procedures that are particularly illuminating in that they highlight both the significance of error and the benefits of the intentional constructive use of the concept of error in scientific investigation.

## References

Alder, K.: 2002, The Measure of All Things: The Seven-Year Odyssey and Hidden Error That Transformed the World, The Free Press, New York.

Bacon, F.: 2000 (originally 1620), The New Organon, Cambridge University Press, Cambridge.

Bady, R.: 1979, 'Students' Understanding of the Logic of Hypothesis Testing', Journal of Research in Science Teaching 16(1), 61-65.

Cochran, W.G.: 1968, 'The Design of Experiments', in International Encyclopedia of the Social Sciences 5, MacMillan, New York, 1968, 245-254.

Confrey, J.: 1990, 'A Review of the Research on Student Conceptions in Mathematics, Science, and Programming', in C. Cazden (ed.), Review of Research in Education 16, American Educational Research Association, Washington, D.C., 3-56.

Fairbrother, R., Hackling, M., & Cowan, E.: 1997, 'Is This the Right Answer?', International Journal of Science Education 19(8), 887-894.

Feynman, R.: 1965, The Character of Natural Law, The MIT Press, Cambridge, MA.

Fisher, K. & Lipson, J.: 1986, 'Twenty Questions About Student Error', Journal of Research in Science Teaching 23(9), 783-203.

Fisher, R.A.: 1949 (originally1935), The Design of Experiments, Oliver & Boyd, Edinburgh.

Goethe, J.W. von: 1988 (originally 1792), 'The Experiment as Mediator Between Subject and Object', in D. Miller (ed.), Goethe: Scientific Studies, Suhrkamp Publishers, New York.

Kaplan, A.: 1964, The Conduct of Inquiry: Methodology for Behavioral Science, Chandler Publishing Company, Scranton, PA.

Kuhn, D., Amsel, E., & O'Loughlin M.: 1988, The Development of Scientific Thinking Skills, Academic Press, San Diego, CA.

Lubben, F. & Millar, R.: 1996, 'Children's Ideas About the Reliability of Experimental Data', International Journal of Science Education 18(8), 955-968.

Mach, E.: 1976 (originally1905), Knowledge and Error: Sketches on the Psychology of Inquiry, D. Reidel Publishing Company, Dordrecht.

Martin, M.: 1970, 'Anomaly-Recognition and Research in Science Education', Journal of Research in Science Teaching 7, 187-190.

Mayo, D.: 1996, Error and the Growth of Experimental Knowledge, The University of Chicago Press, Chicago.

Mosteller, F.: 1968, 'Nonsampling Errors', in International Encyclopedia of the Social Sciences 5, MacMillan, New York, 113-131.

Ministry of Education & Training, Ontario, Canada: 1999, Ontario Science Curriculum – 9th & 10th Grade, found at: http://www.edu.gov.on.ca/eng/document/curricul/secondary/science/scieful.html.

Novak, J.D. (ed.): 1987, Proceedings of the Second International Seminar: Misconceptions and Educational Strategies in Science and Mathematics 1-2, Cornell University, Department of Education, Ithaca, NY.

Park, J., Kim, I., Kim, M., & Lee, M.: 2001, 'Analysis of Student's Processes of Confirmation and Falsification of Their Prior Ideas About Electrostatics', International Journal of Science Education 23(12), 1219-1236.

Piaget J. & Inhelder B.: 1975, The Origin of the Idea of Chance in Children, W.W. Norton & Co., New York.

Polanyi, M.: 1958, Personal Knowledge: Towards a Post-Critical Philosophy, University of Chicago Press, Chicago.

Popper, K.: 2002 (originally1963), Conjectures and Refutations, Routledge, London.

Roschelle, J.: 1998. 'Beyond Romantic Versus Sceptic: A Microanalysis of Conceptual Change in Kinematics', International Journal of Science Education 20(9), 1025-1042.

Rosenbaum, P.R.: 1995, Observational Studies, Springer-Verlag, New York.

Séré, M., Journeaux, R., & Larcher, C.: 1993, 'Learning the Statistical Analysis of Measurement Errors', International Journal of Science Education 15(4), 427-438.

Taylor, J.: 1982, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, University Science Books, Oxford University Press, Mill Valley, CA.

Zachos, P.: 2003 (in press), 'Pendulum Phenomena and the Assessment of Scientific Inquiry Capabilities', Science and Education (to be published in a special issue devoted to the International Pendulum Project).

Zachos, P., Hick, T., Doane, W., & Sargent, C.: 2000, 'Setting Theoretical and Empirical Foundations for Assessing Scientific Inquiry and Discovery in Educational Programs', Journal of_Research in Science Teaching 37(9), 938-962.